A HEIGHT GAP THEOREM FOR FINITE SUBSETS OF $SL_n(\overline{\mathbb{Q}})$ AND NON AMENABLE SUBGROUPS

EMMANUEL BREUILLARD

Abstract. We show a global adelic analog of the classical Margulis Lemma from hyperbolic geometry. We introduce a conjugation invariant normalized height $\widehat{h}(F)$ of a finite set of matrices F in $SL_n(\overline{\mathbb{Q}})$ which is the adelic analog of the minimal displacement on a symmetric space. We then show, making use of theorems of Bilu and Zhang on the equidistribution of Galois orbits of small points, that $\widehat{h}(F) > \varepsilon$ as soon as F generates a non-virtually solvable subgroup of $SL_n(\overline{\mathbb{Q}})$, where $\varepsilon = \varepsilon(n) > 0$ is an absolute constant.

1. Introduction

Definitions.

In this paper we will be concerned with the geometric and arithmetic behavior of power sets $F^n = F \cdot ... \cdot F$ for $n \in \mathbb{N}$, where F is a finite subset of $SL_d(\overline{\mathbb{Q}})$. To study those, we introduce the quantity $\widehat{h}(F)$, which we call normalized height of F and study its properties. It is an invariant of the diagonal action by conjugation of SL_d on SL_d^k, where $k = Card(F)$, and it is a measure of the combined *spectral radius* of F (i.e. the rate of exponential growth of F^n) at all places v, where v varies among all possible equivalence classes of non trivial absolute values on the number field of matrix coefficients of F. Before going further, let us give some definitions.

Let $d \geq 1$ be an integer, $\overline{\mathbb{Q}}$ be the field of algebraic numbers, and $K \leq \overline{\mathbb{Q}}$ a number field. We let V_K be the set of equivalence classes of absolute values on K and $n_v = [K_v : \mathbb{Q}_p]$ the degree of the completion K_v of K over the closure $\mathbb{Q}_p$ of $\mathbb{Q}$ in K_v. We normalise the absolute value $|\cdot|_v$ on K_v so that its restriction to $\mathbb{Q}_p$ is the standard absolute value, i.e. $|p|_v = \frac{1}{p}$. To any finite subset F of square matrices in $M_d(K)\backslash\{0\}$ we associate the following **height**

$$h(F) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ ||F||_v$$

where $\log^+ = \max\{0, \log\}$ and $||F||_v = \max\{||f||_v, f \in F\}$. Here $||f||_v$ is the operator norm on $M_d(K_v)$ associated to the *standard norm* on K_v^d. We define

Date: January 2008.

the standard norm for $x \in K_v^d$ to be the sup norm $||x||_v = \max_{1 \leq i \leq d} |x_i|_v$ if v is ultrametric and the Euclidean norm $||x||_v = \sqrt{\sum_{i=1}^d |x_i|_v^2}$ otherwise. If $d = 1$, this notion coincides with the Weil height of an algebraic number (see e.g. [3]).

We can now define the **normalized height** $\widehat{h}(F)$ as

$$\widehat{h}(F) = \lim_{n \to +\infty} \frac{1}{n} h(F^n)$$

This limit exists by subadditivity of the height. Unlike $h(F)$, $\widehat{h}(F)$ is independent of the choice of norms $||x||_v$ on K_v^d used to define it. Another way to describe $\widehat{h}(F)$ is in terms of spectral radius (see Section 2.2 below) ; for instance if $F = \{A\}$ is a singleton, then $\widehat{h}(F) = h([1, \lambda_1, ..., \lambda_d])$, where $(\lambda_1, ..., \lambda_d)$ are the eigenvalues of A and $h([1, \lambda_1, ..., \lambda_d])$ the standard Weil height of the point $[1, \lambda_1, ..., \lambda_d]$ in the projective space $\mathbb{P}^d(\overline{\mathbb{Q}})$.

The normalized height enjoys the following simple relation $\widehat{h}(F^n) = n \cdot \widehat{h}(F)$ for $n \in \mathbb{N}$. A finite set F satisfies $\widehat{h}(F) = 0$ if and only if F generates a quasi-unipotent subgroup, i.e. a group all of whose elements have only roots of unity as eigenvalues (Proposition 3.2). Moreover, provided F is in "generic position", $\widehat{h}(F)$ is comparable, up to a multiplicative constant to the infimum $\inf h(gFg^{-1})$ of the heights of conjugates of F by elements $g \in GL_d(\overline{\mathbb{Q}})$ (see Proposition 3.3 below).

Height gap.

The main result of this paper establishes the existence of a uniform gap for the normalized height of subsets F generating a non amenable subgroup of $SL_d(\overline{\mathbb{Q}})$. We have:

Theorem 1.1. There is a constant $\varepsilon = \varepsilon(d) > 0$ such that if F is a finite subset of $SL_d(\overline{\mathbb{Q}})$ generating a non amenable subgroup, then $\widehat{h}(F) > \varepsilon$.

The constant $\varepsilon(d)$ can be made explicit in principle, although we make no attempt here to give lower bound (see Remark 2.4).

Recall that, as follows for instance from the Tits alternative ([27]), amenable subgroups of $SL_d(\overline{\mathbb{Q}})$ are precisely the virtually solvable subgroups, i.e. those subgroups which contain a solvable subgroup of finite index. As it turns out, for each integer $k \geq 2$, the set of k-tuples F in $SL_d(\overline{\mathbb{Q}})$ which generate a virtually solvable subgroup forms a closed algebraic subvariety of $SL_d(\overline{\mathbb{Q}})^k$. Therefore Theorem 1.1 implies that the set of points with small normalized height in $SL_d(\overline{\mathbb{Q}})^k$ is not Zariski-dense. This is reminiscent of the Bogomolov conjecture proved by Ullmo and Zhang (see [29], [33], [25]), which

asserts that, given an abelian variety, the set of points with small Néron-Tate height on an algebraic subvariety which is not a finite union of torsion cosets of abelian subvarieties is not Zariski-dense.

In a subsequent paper [8] (see also [7]) we will show how to use Theorem 1.1 in order to obtain a strengthening of the classical Tits alternative. This was in fact our original motivation for Theorem 1.1. In the same vein, but much more straightforwardly, one obtains the following corollary, which answers a question from [1] and is a strengthening of a well known theorem of Shur asserting that finitely generated linear torsion groups are finite.

Corollary 1.2. There is an integer $N = N(d) \in \mathbb{N}$ such that if F is a finite subset of $SL_d(\mathbb{C})$ which generates an infinite subgroup, then $(F \cup F^{-1})^N$ contains an element of infinite order.

The interpretation of $\widehat{h}(F)$ in terms of spectral radius allows to derive the following:

Corollary 1.3. There are constants $\eta = \eta(d) > 0$ and $N_1 = N_1(d) \in \mathbb{N}$ such that if F is a finite subset of $SL_d(\overline{\mathbb{Q}})$ generating a non amenable subgroup, then there is a matrix $w \in F^{N_1}$ with an eigenvalue λ such that $h(\lambda) > \eta$.

In other words Theorem 1.1 allows to construct a short (positive) word w with letters in F which has an eigenvalue of large height. The length of the word is bounded by an absolute constant $N_1 = N_1(d)$. This type of result is crucial in order to build the so-called proximal elements which are needed in various situations, in particular in the Tits alternative.

In the same vein we have:

Corollary 1.4. There is a constant $N_2 = N_2(d) \in \mathbb{N}$, such that if F is a finite subset of $SL_d(\mathbb{C})$ which generates a non amenable subgroup, then there is a matrix $w \in F^{N_1}$ with an eigenvalue λ with the following property. Either there exists an ultrametric absolute value on $\mathbb{Q}(\lambda)$ such that $|\lambda| > 1$, or there is a field homomorphism $\sigma : \mathbb{Q}(\lambda) \hookrightarrow \mathbb{C}$ such that $|\sigma(\lambda)| \geq 2$.

Geometric Interpretation and the Margulis Lemma.

Theorem 1.1 has also the following geometric interpretation. Recall that the classical Margulis Lemma (see [28]) asserts that if $S = \mathbb{H}^n$ is the hyperbolic n-space, or more generally any real symmetric space of non compact type endowed with its Riemannian metric d, then there is a positive constant $\varepsilon = \varepsilon(S) > 0$ such that the following holds: suppose F is a finite set of isometries of S such that $\max_{f \in F} d(f \cdot x, x) < \varepsilon$ for some point $x \in S$ and suppose F lies in a discrete subgroup of isometries of S, then F generates a virtually nilpotent subgroup. This lemma has several important consequences for the geometry and topology of hyperbolic manifolds and locally

symmetric spaces, such as the structure of cusps and the thick-thin decomposition ([28]), or lower bounds for the covolume of lattices in semisimple Lie groups (see [30], [14]).

What happens if one removes the discreteness assumption on the group generated by F and assumes instead that F consists of elements which are rational over some number field K ? Of course the Margulis Lemma no longer holds as such, in particular because $\varepsilon(S)$ tends to 0 as $\dim S$ tends to infinity. However Theorem 1.1 gives a kind of substitute. As will be shown below (see Section 2.2) the normalized height $\widehat{h}(F)$ is always bounded above by the quantity $e(F)$, which we call minimal height, and which encodes, as a weighted sum over all places $v \in V_K$, the minimal displacement of F on each symmetric space or Bruhat-Tits building X_v associated to $SL_d(K_v)$. In particular the height gap $\widehat{h}(F) > \varepsilon$ obtained in Theorem 1.1 implies there always is a natural space X_v (symmetric space or Bruhat-Tits building of SL_d) where F acts with a large displacement. More precisely:

Corollary 1.5. Let $d \in \mathbb{N}$ and for a local field k denote by X_k the symmetric space or Bruhat-Tits building of $SL_d(k)$. Then there is a constant $\varepsilon = \varepsilon(d) > 0$ with the following property. Let K be a number field and F a finite subset of $SL_d(K)$ which generates a non amenable subgroup Γ, then either for some finite place v of K, the subgroup Γ acts (simplicially) without global fixed point on the Bruhat-Tits building X_{K_v}, or for some embedding $\sigma : K \hookrightarrow \mathbb{C}$

$$\inf_{x \in X_{\mathbb{C}}} \max_{f \in F} d(\sigma(f) \cdot x, x) > \varepsilon$$

where $d(\cdot, \cdot)$ is the left invariant Riemannian metric on $X_{\mathbb{C}}$.

The crucial point here of course is that ε is independent of the number field K. Thus Theorem 1.1 can be seen as a uniform Margulis Lemma for all S-arithmetic lattices of a given Lie type. For example, it is uniform over all $SL_2(\mathcal{O}_K)$ where K can vary among all number fields, even though those groups can be lattices of arbitrarily large rank.

Outline of the proof of Theorem 1.1.

The first part of the proof consists in reducing to the situation when F is a 2-element set $F = \{A, B\}$, where A and B are two regular semisimple elements in an absolutely almost simple algebraic group $\mathbb{G}$ of adjoint type and F generates a Zariski-dense subgroup of $\mathbb{G}$. It is not hard to see that the existence of a gap for $\widehat{h}(F)$ when computed in the adjoint representation of $\mathbb{G}$ implies the existence of a gap for $\widehat{h}(F)$ when computed in any finite dimensional linear representation of $\mathbb{G}$. We thus reduce to the adjoint representation of $\mathbb{G}$. The reduction from an arbitrary finite set F to a 2-element

set makes use of a lemma due to Eskin-Mozes-Oh [13] ("escaping subvarieties" Lemma 4.16), which allows, given any non trivial algebraic relation between pairs $\{x, y\}$ of elements in $\mathbb{G}$ to find two short words in $\{x, y\}$ which no longer satisfy this relation. This lemma is also used later on and is an essential tool here.

As we mentioned above, one may interpret $\widehat{h}(F)$ in terms of the combined minimal displacement $e(F)$ of F on all symmetric spaces and Bruhat-Tits buildings that arise through the various completions of the number field. The quantity $e(F)$ is defined as the weighted sum of the logarithm of the minimal norms $E_v(F) = \inf\{||gFg^{-1}||_v, g \in GL_d(\overline{K_v})\}$. Crucial to this correspondence is a spectral radius formula for sets of matrices (Lemma 2.1 below) which compares the minimal displacement of F (or equivalently $E_v(F)$) with the minimal displacement of each individual matrix in the power set F^{d^2} (or equivalently its maximal eigenvalue). As a consequence, $\widehat{h}(F)$ is small if and only if $e(F)$ is small.

In the second part of the proof, we fix a place v and work in $\mathbb{G}(K_v)$. Given A, B in $\mathbb{G}(K_v)$, with A in a maximal torus T of $\mathbb{G}(K_v)$, we obtain local estimates for the minimal displacement of the action of B restricted to the maximal flat associated to T. These estimates are obtained via the Iwasawa decomposition working our way through all positive roots of A starting from the maximal one. At the end we get an upper bound for $\inf_{t_v \in T} ||t_v B t_v^{-1}||_v$ which involves $E_v(F)$ on the one hand and the gap $|1 - \alpha(A)|_v$ between the roots of $\alpha(A)$ and 1 on the other hand.

In the last part of the proof, we put all our local estimates together and make crucial use of the product formula, so as to obtain an upper bound for the weighted sum of all $\inf_{t_v \in T} \log ||t_v B t_v^{-1}||_v$ in terms of $e(F)$ and the average of the $\log |1 - \alpha(A)|_v$ over all archimedean places v, for each root α. When $e(F)$ is small this upper bound becomes also small. Indeed, since the height of each $\alpha(A)$ is small, we can invoke Bilu's equidistribution theorem : the Galois conjugates of $\alpha(A)$ equidistribute on the unit circle ([2]). Hence the average of the $\log |1 - \alpha(A)|_v$'s gives a negligible contribution.

Finally, considering a suitably chosen T-invariant regular map f on $\mathbb{G}$ (a suitable matrix coefficient of B will do), we use the above upper bound to show that the height of $f(B)$ as well as $f(B^i)$ for larger and larger $i \in \mathbb{N}$, becomes small when $e(F)$ is small. However, by a theorem of Zhang [32] on small points of algebraic tori, this must force a non trivial algebraic relation between the $f(B^i)$'s. Finally the Eskin-Mozes-Oh lemma quoted above provides the desired contradiction, as we may have chosen $F = \{A, B\}$ to avoid this relation to begin with.

Contents

1. Introduction 1
2. Minimal height and displacement 6
3. Statement of the results 13
4. Preliminary reductions 15
5. Local estimates on Chevalley groups 24
6. Global bounds on arithmetic heights 31
7. Proof of the statements of Section 3 36
References 44

2. Minimal height and displacement

2.1. **Local notions of minimal norm, spectral radius and minimal displacement.** Let k be a local field of characteristic 0. Let $\|\cdot\|_k$ be the standard norm on k^d, that is the canonical Euclidean (resp. Hermitian) norm if $k = \mathbb{R}$ (resp. $\mathbb{C}$) and the sup norm ($\|x\|_k = \max_i |x_i|_k$) if k is non Archimedean. We will also denote by $\|\cdot\|_k$ the operator norm induced on the space of d by d matrices $M_d(k)$ by the standard norm $\|\cdot\|_k$ on k^d. Let Q be a bounded subset of matrices in $M_d(k)$. We set

$$\|Q\|_k = \sup_{g\in Q} \|g\|_k$$

and call it the *norm of* Q. Let $\overline{k}$ be an algebraic closure of k. It is well known (see Lang's Algebra [18]) that the absolute value on k extends to a unique absolute value on $\overline{k}$, hence the norm $\|\cdot\|_k$ also extends in a natural way to $\overline{k}^d$ and to $M_d(\overline{k})$. This allows to define the *minimal norm* of a bounded subset Q of $M_d(k)$ as

$$E_k(Q) = \inf_{x\in GL_d(\overline{k})} \left\|xQx^{-1}\right\|_k$$

We will also need to consider the *maximal eigenvalue of* Q, namely

$$\Lambda_k(Q) = \max\{|\lambda|_k,\ \lambda \in spec(q), q \in Q\}$$

where $spec(q)$ denotes the set of eigenvalues (the spectrum) of q in $\overline{k}$. We also set $Q^n = Q \cdot ... \cdot Q$ be the set of all products of n elements from Q. Finally, we introduce the *spectral radius* of Q, that is

$$R_k(Q) = \lim_{n\to+\infty} \|Q^n\|_k^{\frac{1}{n}}$$

in which the limit exists (and coincides with $\inf_{n\in\mathbb{N}} \|Q^n\|_k^{\frac{1}{n}}$) because the sequence $\{\|Q^n\|_k\}_n$ is sub-multiplicative.

These quantities are related to one another. The key property concerning them is given in the following proposition, a weaker version of which was proven in [6]. The intuition behind this result was inspired by the work of Eskin-Mozes-Oh [13], where a result of a similar nature appears inside their argument.

Lemma 2.1. (Spectral Radius Formula for Q) Let Q be a bounded subset of $M_d(k)$.

(a) if k is non Archimedean, there is an integer $q \in [1, d^2]$ such that $\Lambda_k(Q^q) = E_k(Q)^q$.

(b) if k is Archimedean, there is a constant $c = c(d) \in (0, 1)$ independent of Q and an integer $q \in [1, d^2]$ such that $\Lambda_k(Q^q) \geq c^q \cdot E_k(Q)^q$.

Proof. Let K be a field. We make use of two well-known theorems. The first is a theorem of Wedderburn (see Curtis-Reiner [11] 27.27) that if an algebra A over K has a linear basis over K consisting of nilpotent elements, then $A^m = 0$ for some integer m. The second is a theorem of Engel (see Jacobson [16]) that if A is a subset of $M_d(K)$ such that $A^m = 0$ for some integer m, then A can be simultaneously conjugated in $GL_d(K)$ inside $N_d(K)$, the subalgebra of upper triangular matrices with zeroes on and below the diagonal. Combined together, these facts yield:

Lemma 2.2. If Q is any subset of $M_d(K)$ such that Q^q contains only nilpotent matrices for every q, $1 \leq q \leq d^2$, then there is $g \in GL_d(K)$ such that $gQg^{-1} \subset N_d(K)$.

Proof. Since $\dim_K M_d(K) \leq d^2$, the K-algebra generated by Q has a linear basis made of elements in $\cup_{1 \leq q \leq d^2} Q^q$. By Wedderburn and Engel, the result follows. □

We first quickly prove (b). We argue by contradiction. There is a sequence Q_n with $E(Q_n) = 1$ while $\max_{1 \leq q \leq d^2} \Lambda(Q_n^q)^{\frac{1}{q}}$ tends to 0. Up to conjugating by some $g_n \in GL_d(\mathbb{C})$, we may assume that $||Q_n|| \leq 1 + \frac{1}{n}$, and passing to a Haussdorff limit, we obtain a compact set Q with $E(Q) = ||Q|| = 1$, while $\max_{1 \leq q \leq d^2} \Lambda(Q^q)^{\frac{1}{q}} = 0$. But this is a contradiction with lemma 2.2 as $E(N(\mathbb{C})) = 0$. This proves (b).

In order to prove (a) we first show:

Lemma 2.3. Let $d \in \mathbb{N}$. There exists an integer $N = N(d) \in \mathbb{N}$ with the following property. Let k be a non archimedean local field with absolute value $|\cdot|_k$ and $\mathcal{O}_k$ its ring of integers. Let Q be a subset of $M_d(\mathcal{O}_k)$ such that for each integer $q \in [1, d^2]$ every element of Q^q has all its eigenvalues of absolute value at most $|\pi|_k^N$, where π is a uniformizer for $\mathcal{O}_k$. Then there is $g \in GL_d(k)$ such that gQg^{-1} belongs to $\pi M_d(\mathcal{O}_k)$.

Proof. We argue by contradiction. This means that we have a sequence of local fields k_n and subsets Q_n in $M_d(\mathcal{O}_{k_n})$ such that $||gQ_ng^{-1}||_{k_n} \geq 1$ for all $g \in GL_d(k_n)$ and all eigenvalues of Q_n^q have absolute value at most $|\pi_n|_{k_n}^n$. Let us consider a non-principal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and form the ultraproduct $A = \prod_{\mathcal{U}} \mathcal{O}_{k_n}$. First let us decide that we have chosen the absolute value $|\cdot|_n$ on k_n in such a way that $|\pi_n|_n = \frac{1}{2}$ for every n where π_n is a fixed uniformizer in $\mathcal{O}_{k_n}$. For every $x_n \in \mathcal{O}_{k_n}$ the quantity $|x_n|_n$ may only take values among $2^{-(\mathbb{N}\cup\{\infty\})}$. It follows that for every $x \in A$ represented by $(x_n)_{n\in\mathbb{N}}$, the quantity $|x| := \lim_{\mathcal{U}} |x_n|_n$, which is well defined, may only take values in $2^{-(\mathbb{N}\cup\{\infty\})}$. Let $I = \{x \in A, |x| = 0\}$. Then we check that I is a prime ideal of A, this follows from the standard properties of ultrafilters. Let $\mathcal{O} = A/I$. Similarly we check that $\mathcal{O}$ is a discrete valuation ring with uniformizer π equal to the class of $(\pi_n)_{n\in\mathbb{N}}$ in A/I and that $|\cdot|$ is a well defined absolute value. Let K be the field of fractions of $\mathcal{O}$. It is a field with a non archimedean absolute value and $\mathcal{O} = \{x \in K, |x| \leq 1\}$. Let Q be the class of $(Q_n)_{n\in\mathbb{N}}$ in $M_d(\mathcal{O})$. Then Q^q is the class of $(Q_n^q)_{n\in\mathbb{N}}$ for each q. But by assumption $|a|_n \leq \frac{1}{2^n}$ for every non maximal coefficient a of the characteristic polynomial of any matrix in Q_n^q. It follows that Q^q is made of nilpotent matrices for each q, $1 \leq q \leq d^2$. We may thus apply Lemma 2.2 to Q in $M_d(K)$. There is a matrix $g \in GL_d(K)$ such that $gQg^{-1} \subset N_d(K)$. Write $g = \pi^{-L}\overline{g}$ where $\overline{g} \in M_d(\mathcal{O})$. There is $\overline{\overline{g}} \in M_d(\mathcal{O})$ such that $\overline{g}\overline{\overline{g}} = \det \overline{g}$ which is the transpose of the matrix of minors. We get thus have $\overline{g}Q\overline{\overline{g}} \subset N_d(\mathcal{O})$. This means that there is a function $f(n)$ going to $+\infty$ with n such that $\overline{g}_nQ_n\overline{\overline{g}}_n \subset N_d(\mathcal{O}_{k_n})$ $\mathrm{mod}\, \pi_n^{f(n)}$ for most n's (i.e. for a set of n's belonging to $\mathcal{U}$). In particular for every $M \in \mathbb{N}$, for most n's one may find a diagonal matrix $h_n \in GL_d(k_n)$ such that $h_n\overline{g}_nQ_n\overline{\overline{g}}_nh_n^{-1} \subset \pi_n^{M+1}M_d(\mathcal{O}_{k_n})$. Finally note that $\det \overline{g} \in \mathcal{O}\backslash\{0\}$ so that if $(\overline{g}_n)_n$ is a representative of $\overline{g}$ in $M_d(A)$, there is $M \in \mathbb{N}$ such that $|\det \overline{g}_n|_n \geq 2^{-M}$ for most $n \in \mathbb{N}$. Hence $h_n\overline{g}_nQ_n\overline{g}_n^{-1}h_n^{-1} \subset \pi_nM_d(\mathcal{O}_{k_n})$ for most n's, which is the desired contradiction. □

We can now prove (a). Let π a uniformizer for k and let $\delta \geq 0$ be such that $\max_{1\leq q\leq d^2} \Lambda_k(Q^q)^{\frac{1}{q}} = |\pi|_k^\delta E_k(Q)$. Assume by contradiction that $\delta > 0$. Let $m \geq N(d)/\delta$. Let $k_1 = k(\pi_1)$ where $\pi_1^m = \pi$ and $F_{k_1}(Q) = \min_{x\in GL_d(k_1)} \|xQx^{-1}\|_{k_1}$. Up to conjugating by $x \in GL_d(k_1)$, we may assume that $F_{k_1}(Q) = ||Q||_{k_1} \geq E_k(Q)$. Let $Q_0 = \frac{Q}{||Q||_{k_1}}$. Then

$$\max_{1\leq q\leq d^2} \Lambda_k(Q_0^q)^{\frac{1}{q}} \leq |\pi_1|_{k_1}^{\delta m} \leq |\pi_1|_{k_1}^{N(d)}$$

while $F_{k_1}(Q_0) = 1$. But this obviously contradicts Lemma 2.3. This ends the proof of (a).

□

Remark 2.4. The proof of (b) was by contradiction and gave no indication as how large c is. This is the only place in this paper (and hence in the determination of the height gap $\varepsilon(d)$ from Theorem 1.1) where we have a constant which is not explicitable in principle. In a subsequent paper we will provide another proof of (b) which is effective and gives a lower bound of order $\exp(-d^{d^2})$ for $c(d)$.

This allows to explain the relationships between minimal norm, spectral radius and maximal eigenvalue:

Proposition 2.5. Let Q be a bounded subset of $M_d(k)$ containing 1. We have

(i) $\Lambda_k(Q) \leq E_k(Q) \leq \|Q\|_k$, and $R_k(gQg^{-1}) = R_k(Q)$ for any $g \in GL_d(\overline{k})$,

(ii) $\Lambda_k(Q^n) \geq \Lambda_k(Q)^n$, $E_k(Q^n) \leq E_k(Q)^n$ and $R_k(Q^n) = R_k(Q)^n$ for every $n \in \mathbb{N}$,

(iii) $R_k(Q) = \lim_{n\to+\infty} E_k(Q^n)^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} E_k(Q^n)^{\frac{1}{n}}$,

(iv) $R_k(Q) = \lim_{n\to+\infty} \Lambda_k(Q^n)^{\frac{1}{n}} = \sup_{n\in\mathbb{N}} \Lambda_k(Q^n)^{\frac{1}{n}}$,

(v) if k is non Archimedean, $R_k(Q) = E_k(Q)$,

(vi) if k is Archimedean, $c \cdot E_k(Q) \leq R_k(Q) \leq E_k(Q)$, where c is the constant from Lemma 2.1 (b).

Proof. Items (i) and (ii) are clear from the definitions. Let us first show (iii). We have $E_k(Q^n) \leq ||Q^n||_k$ for every $n \in \mathbb{N}$, hence $\limsup E_k(Q^n)^{\frac{1}{n}} \leq R_k(Q)$. On the other hand, $R_k(Q) = R_k(gQg^{-1}) \leq ||gQg^{-1}||_k$ for every $g \in GL_d(\overline{k})$. Hence $R_k(Q) \leq E_k(Q)$ and for every $n \in \mathbb{N}$, $R_k(Q)^n = R_k(Q^n) \leq E_k(Q^n)$, hence $R_k(Q) \leq \liminf E_k(Q^n)^{\frac{1}{n}}$. So we have shown that $\lim E_k(Q^n)^{\frac{1}{n}}$ exists and equals $R_k(Q)$. Furthermore, for every $n, p \in \mathbb{N}$, $E_k(Q^{np})^{\frac{1}{np}} \leq E_k(Q^p)^{\frac{1}{p}}$. Letting n tend to $+\infty$, we obtain $R_k(Q) \leq E_k(Q^p)^{\frac{1}{p}}$. Hence $R_k(Q) = \inf_{n\in\mathbb{N}} E_k(Q^n)^{\frac{1}{n}}$.

Now consider (iv). It is clear that as $\Lambda_k(Q^n) \leq E_k(Q^n)$, we have $\limsup \Lambda_k(Q^n)^{\frac{1}{n}} \leq R_k(Q)$. On the other hand, writing $n = mq + k$ for any $n \in \mathbb{N}$, with $0 \leq k < q$, we get from Lemma 2.1, $\Lambda_k(Q^n)^{\frac{1}{n}} \geq c^{\frac{q}{n}} \cdot E_k(Q^m)^{\frac{q}{mq+k}}$ (where $c = 1$ if k is non Archimedean) which forces $\liminf \Lambda_k(Q^n)^{\frac{1}{n}} \geq R_k(Q)$. Hence finally $\lim_{n\to+\infty} \Lambda_k(Q^n)^{\frac{1}{n}}$ exists and equals $R_k(Q)$. Since for every $n, p \in \mathbb{N}$ we have $\Lambda_k(Q^{np}) \geq \Lambda_k(Q^p)^n$, by letting n tend to $+\infty$, we indeed get $R_k(Q) = \sup_{p\in\mathbb{N}} \Lambda_k(Q^p)^{\frac{1}{p}}$.

Now (v). From (iii) and (iv) we clearly have for any $q \in \mathbb{N}$ $\Lambda_k(Q^q)^{\frac{1}{q}} \leq R_k(Q) \leq E_k(Q)$. If k is non Archimedean, then this combined with Lemma 2.1 (a) shows the desired identity. If k is Archimedean, then it gives $\Lambda_k(Q^q) \leq R_k(Q)^q$, which when combined with Lemma 2.1 (b) gives $c \cdot E_k(Q) \leq R_k(Q)$. □

Remark 2.6. Observe that when $k = \mathbb{R}$ or $\mathbb{C}$, then we may have $R_k(Q) < E_k(Q)$. For instance, consider $Q = \{1, T, S\} \subset SL_2(\mathbb{Z})$, where T and S are the matrices corresponding to the standard generators of $PGL_2(\mathbb{Z})$, i.e. T acts by translation by 1 and S by inversion around the circle of radius 1 in the upper-half plane. Then it is easy to compute $E_k(Q) = \sqrt{2} = ||tQt^{-1}||_k$ where t is the diagonal matrix $t = diag(\sqrt[4]{2}, \frac{1}{\sqrt[4]{2}})$. On the other hand, one can check that $||tQ^2t^{-1}||_k < 2$, so $R_k(Q) \leq E_k(Q^2)^{\frac{1}{2}} < E_k(Q)$.

Note that if Q belongs to $SL_d(k)$, then $E_k(Q) \geq R_k(Q) \geq \Lambda_k(Q) \geq 1$. The following lemma explain what happens if these quantities are close or equal to 1.

Lemma 2.7. Suppose k is Archimedean (i.e. $k = \mathbb{R}$ or $\mathbb{C}$). Then for every $\varepsilon > 0$ and $T > 0$ there is $N_3 = N_3(d, \varepsilon, T) \in \mathbb{N}$ such that if Q is a bounded subset of $SL_d(k)$ such that $E_k(Q) > 1 + \varepsilon$, then $\Lambda_k(Q^q) \geq T$ for some $q \in \mathbb{N}$, $1 \leq q \leq N_3$. In particular, there is $\delta = \delta(d, \varepsilon) > 0$ such that if $E_k(Q) > 1+\varepsilon$, then $R_k(Q) > 1 + \delta$. Moreover $E_k(Q) = 1$ iff $R_k(Q) = 1$.

The lemma follows readily from the following result:

Proposition 2.8. *Suppose k is Archimedean (i.e. $k = \mathbb{R}$ or $\mathbb{C}$). Then for every $n \in \mathbb{N}$ and every bounded subset Q of $SL_d(k)$ containing 1, we have*

$$E_k(Q^n) \geq E_k(Q)^{\sqrt{\frac{n}{8d}}} \tag{1}$$

And

$$\log R_k(Q) \geq c_1 \cdot \log E_k(Q) \cdot \min\{1, \log E_k(Q)\} \tag{2}$$

where $c_1 = c_1(d) > 0$ is a positive constant.

Proof. Let $r_n = \inf_{x \in X} \max_{g \in Q^n} d(gx, x)$ where (X, d) is the symmetric space associated to $SL_d(k)$. We claim that $r_1 \leq \sqrt{\frac{8}{n}} r_n$. Fix n and let ε be arbitrarily small. For each $k \leq n$ let $x_k \in X$ be such that the lower bound r_k is realized up to an error of ε. Choose some $k \leq n$ such that $r_k - r_{k-1} \leq \frac{r_n}{n}$. For each $g \in Q$, the balls of radius $r_k + \varepsilon$ centered at x_k and gx_k both contain $gQ^{k-1}x_k$. By the $CAT(0)$ inequality for the median in X, $gQ^{k-1}x_k$ lies in the ball centered at the mid point between x_k and gx_k and with squared radius at most $(r_k + \varepsilon)^2 - \frac{1}{4}d(x_k, gx_k)^2$. Hence $\frac{1}{4}d(x_k, gx_k)^2 \leq (r_k + \varepsilon)^2 - r_{k-1}^2$. And $r_1^2 \leq 4(\frac{r_n}{n} + \varepsilon)(2r_n + \varepsilon)$. As ε was arbitrary, we do get $r_1 \leq \sqrt{\frac{8}{n}} r_n$.

Now by definition of the metric on X, we have $\log E_k(Q^n) \in [\frac{1}{\sqrt{d}}, 1] r_n$. Hence (1). For (2), note that for every n, by Lemma 2.1 there is $q \leq d^2$ such that $\Lambda_k(Q^{qn})^{\frac{1}{q}} \geq cE_k(Q^n) \geq cE_k(Q)^{\sqrt{\frac{n}{8d}}}$ and hence $R_k(Q) \geq c^{\frac{1}{n}} E_k(Q)^{\sqrt{\frac{1}{8dn}}}$. Optimizing in n we obtain a constant $c_1 = c_1(d)$ for which (2) holds. □

Remark 2.9. Note that for arbitrarily large n one can find finite sets Q containing 1 (in $SL_2(\mathbb{R})$ for instance) such that $\Lambda_k(Q^n) = 1$ while $E_k(Q) > 1$. Indeed consider the set Q consisting of 1 and two elliptic elements fixing two different but extremely close fixed points in the Poincaré disc. This is another way to see that the constant c in Lemma 2.1 (b) cannot be 1.

Remark 2.10. Observe that if $Q \subset SL_d(k)$, then adding the identity to Q does not modify our quantities. Namely if $Q_1 = Q \cup \{Id\}$, then $E_k(Q_1) = E_k(Q)$, $\Lambda_k(Q_1) = \Lambda_k(Q)$ and also $R_k(Q_1) = R_k(Q)$. For the last identity, note that for all $n \in \mathbb{N}$, there is $m \leq n$ such that $\Lambda_k(Q_1^n) = \Lambda_k(Q^m) \leq R_k(Q)^m \leq R(Q)^n$, since $R_k(Q) \geq 1$, hence as n tends to $+\infty$, $R_k(Q_1) \leq R_k(Q)$, while the converse inequality is clear.

2.2. **Height, normalized height and minimal height.** For any rational prime p let us fix an algebraic closure $\overline{\mathbb{Q}_p}$ of the field of p-adic numbers $\mathbb{Q}_p$. We take the standard normalization of the absolute value on $\mathbb{Q}_p$ (i.e. $|p|_p = \frac{1}{p}$). It admits a unique extension to $\overline{\mathbb{Q}_p}$, which we denote by $|\cdot|_p$. Let $\overline{\mathbb{Q}}$ be the field of all algebraic numbers and K a number field. Let V_K be the set of equivalence classes of valuations on K. For $v \in V_K$ let K_v be the corresponding completion. For each $v \in V_K$, K_v is a finite extension of $\mathbb{Q}_p$ for some prime p. We normalize the absolute value on K_v to be the unique one which extends the standard abolute value on $\mathbb{Q}_p$. Namely $|x|_v = |N_{K_v|\mathbb{Q}_p}(x)|_p^{\frac{1}{n_v}}$ where $n_v = [K_v : \mathbb{Q}_p]$. Equivalently K_v has n_v different embeddings in $\overline{\mathbb{Q}_p}$ and each of them gives rise to the same absolute value on K_v. We identify $\overline{K_v}$, the algebraic closure of K_v with $\overline{\mathbb{Q}_p}$. Let V_f be the set of finite places and V_∞ the set of infinite places.

Let $d \in \mathbb{N}$ be an integer $d \geq 2$. Let F be a finite subset in $SL_d(K)$. For $v \in V_K$, in order not to surcharge notation, we will use the subscript v instead of K_v in the quantities $E_v(F) = E_{K_v}(F)$, $\Lambda_v(F) = \Lambda_{K_v}(F)$, etc.

Recall that if $x \in K$ then its height is by definition (see e.g. [3]) the following quantity

$$h(x) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ |x|_v$$

It is well defined (i.e. independent of the choice of $K \ni x$). Let us similarly define the height of a matrix $f \in M_d(K)$ by

$$h(f) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ ||f||_v$$

and the height of a finite set F of matrices in $M_d(K)$ by

$$h(F) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ ||F||_v \tag{3}$$

where $n_v = [K_v : \mathbb{Q}_v]$. We also define the *minimal height* of F as:

$$e(F) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ E_v(F) \tag{4}$$

and the *normalized height* of F as:

$$\widehat{h}(F) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in V_K} n_v \log^+ R_v(F)$$

For any height h, we also set $h = h_\infty + h_f$, where h_∞ is the infinite part of h (i.e. the part of the sum over the infinite places of K) and h_f is the finite part of h (i.e. the part of the sum over the finite places of K). Note that these heights are well defined independently of the number field K such that $F \subset M_d(K)$.

Remark 2.11. If we choose another basis of $\overline{\mathbb{Q}}^d$ the new height $h_{new}(F)$ differs only from the original height by a bounded error. Indeed there are only finitely many places where the new standard norm may differ from the original one. On the other hand $\widehat{h}(F)$ is independent of the choice of basis.

The above terminology is justified by the following facts:

Proposition 2.12. For any finite set F in $M_d(\overline{\mathbb{Q}})$, we have:
(a) $\widehat{h}(F) = \lim_{n\to+\infty} \frac{1}{n} h(F^n) = \inf_{n\in\mathbb{N}} \frac{1}{n} h(F^n)$,
(b) $e_f(F) = \widehat{h}_f(F)$ and $e(F) + \log c \leq \widehat{h}(F) \leq e(F)$ where c is the constant in Lemma 2.1 (b),
(c) $\widehat{h}(F^n) = n \cdot \widehat{h}(F)$ and $\widehat{h}(F \cup \{Id\}) = \widehat{h}(F)$,
(d) $\widehat{h}(xFx^{-1}) = \widehat{h}(F)$ if $x \in GL_d(\overline{\mathbb{Q}})$.

Proof. Since F is finite, there are only finitely many places v such that $||F||_v > 1$. For each such place, $\frac{1}{n} \log^+ ||F^n||_v \to \log^+ R_v(F)$, hence $\frac{1}{n} h(F^n) \to \widehat{h}(F)$. By Prop. 2.5 (vii) we have $E_v(F) = R_v(F)$ if $v \in V_f$, hence $e_f(F) = \widehat{h}_f(F)$, while $c \cdot E_v(F) \leq R_v(F) \leq E_v(F)$ if $v \in V_\infty$, hence $e_\infty(F) + \log c \leq \widehat{h}_\infty(F) \leq e_\infty(F)$. Finally by Prop. 2.5 (ii) $R_v(F^n) = R_v(F)^n$ for every $n \in \mathbb{N}$ and ever place v. Hence $\widehat{h}(F^n) = n \cdot \widehat{h}(F)$. □

We also record the following simple observation:

Proposition 2.13. (a) Note that $e(xFx^{-1}) = e(F)$ for all F finite in $M_d(\overline{\mathbb{Q}})$ and $x \in GL_d(\overline{\mathbb{Q}})$.

(b) $e(F^n) \leq n \cdot e(F)$,

(c) If λ is an eigenvalue of an element of F, then $h(\lambda) \leq e(F)$,

(d) If F is a subset of $SL_d(\overline{\mathbb{Q}})$, then $e(F \cup F^{-1}) \leq (d-1) \cdot e(F)$ and $e(F \cup \{1\}) = e(F)$.

Proof. The first three items are clear. For the last simply observe that $||x^{-1}||_v \leq ||x||_v^{d-1}$ for any $x \in SL_d(K_v)$ as can be seen by expressing those norms in terms of the KAK decomposition of x. □

We can also compare $e(F)$ and $\widehat{h}(F)$ when $\widehat{h}(F)$ is small:

Proposition 2.14. For every $\varepsilon > 0$ there is $\delta = \delta(d, \varepsilon) > 0$ such that if F is a finite set in $SL_d(\overline{\mathbb{Q}})$ with $\widehat{h}(F) < \delta$, then $e(F) < \varepsilon$. Moreover $\widehat{h}(F) = 0$ iff $e(F) = 0$.

This follows immediately from 2.12 b) and the following proposition.

Proposition 2.15. Let c_1 be the constant from Proposition 2.8, then

$$\widehat{h}_\infty(F) \geq \frac{c_1}{4} \cdot e_\infty(F) \cdot \min\{1, e_\infty(F)\}$$

for any finite subset F in $SL_2(\overline{\mathbb{Q}})$.

Proof. From Propositions 2.8 and 2.5 (v), $\widehat{h}_v(F) \geq c_1 \cdot e_v(F) \cdot \min\{1, e_v(F)\}$ for every $v \in V_K$. We may write $e_\infty(F) = \alpha e^+(F) + (1-\alpha)e^-(F)$ where e^+ is the average of the e_v greater than 1 and e^- the average of the e_v smaller than 1. Applying Cauchy-Schwartz, we have $\widehat{h}_v(F) \geq c_1 \cdot (\alpha e^+ + (1-\alpha)(e^-)^2)$. If $\alpha e^+(F) \geq \frac{1}{2} e_\infty(F)$, then $\widehat{h}_v(F) \geq \frac{c_1}{2} e_\infty(F)$, and otherwise $(1-\alpha)e^- \geq \frac{e_\infty}{2}$, hence $\widehat{h}_v(F) \geq (1-\alpha)(e^-)^2 \geq \frac{1}{4} e_\infty^2$. At any case $\widehat{h}_\infty(F) \geq \frac{c_1}{4} \cdot e_\infty(F) \cdot \min\{1, e_\infty(F)\}$. □

3. Statement of the results

We state here our results. The main theorem is the following:

Theorem 3.1. There exists a positive constant $\varepsilon = \varepsilon(d) > 0$ with the following property. If F is a finite subset of $SL_d(\overline{\mathbb{Q}})$ with $\widehat{h}(F) \leq \varepsilon$ then F generates a virtually solvable subgroup of $SL_d(\overline{\mathbb{Q}})$.

Note first that since $\widehat{h}(F)$ is small whenever $e(F)$ is small (according to Proposition 2.14), we may substitute $e(F)$ to $\widehat{h}(F)$ in the theorem.

It is easy to characterize sets of zero normalized height :

Proposition 3.2. If F is a finite subset of $SL_d(\overline{\mathbb{Q}})$, then $\widehat{h}(F) = 0$ if and only if the group generated by F virtually unipotent.

Proof. If $\widehat{h}(F) = 0$, then according to Propositions 2.14 and 2.13, $e(F) = e(F \cup F^{-1}) = 0$, hence $e((F \cup F^{-1})^n) = 0$ for each $n \in \mathbb{N}$. Thus every element from the group $\langle F \rangle$ generated by F has only roots of unity as eigenvalues. However, according to Theorem 6.11 in [22], $\langle F \rangle$ has a finite index subgroup Γ_0 for which no element has a non-trivial root of unity as eigenvalue. Therefore every element in Γ_0 must be unipotent, i.e. Γ_0 is unipotent. Conversely, if $\langle F \rangle$ is virtually unipotent, then every element in $\langle F \rangle$ has its eigenvalues among the roots of unity. In particular, as follows from Proposition 2.5 (iv), $R_v(F) = 1$ for every place v. Hence $\widehat{h}(F) = 0$. □

The above results dealt with small values of the normalized height. The following proposition says in substance that, provided $\langle F \rangle$ acts absolutely irreducibly, the normalized height is attained up to multiplicative and additive constants by the height of some suitable conjugate of F in $SL_d(\overline{\mathbb{Q}})$. We have

Proposition 3.3. If $\mathbb{G} \leq SL_d(\overline{\mathbb{Q}})$ is an absolutely almost simple algebraic subgroup acting irreducibly on $\overline{\mathbb{Q}}^d$, then there is a constant $C > 0$ and a non-empty Zariski-open subset Ω of $\mathbb{G} \times \mathbb{G}$ such that for all $(a, b) \in \Omega$ there exists $g \in \mathbb{G}(\overline{\mathbb{Q}})$ such that for $F = \{a, b\}$

$$h(gFg^{-1}) \leq C \cdot \widehat{h}(F)$$

This proposition is important for applications as it allows to conjugate F back in the "right position" in some sense. Observe that by definition $e(F)$ is equal to the infimum of $h(gFg^{-1})$ when $g = (g_v)_{v \in V_K}$ is allowed to vary among the full group of adèles $GL_d(\mathbb{A})$. This proposition shows that this infinimum is morally attained on principal adèles, i.e. on $GL_d(\overline{\mathbb{Q}})$. The condition that F should belong to a Zariski-open subset is important as easy examples show that the result of the proposition can fail if for instance F fixes a proper subspace.

Remark 3.4. Proposition 3.3 actually holds for a finite set F of arbitrary cardinality with a constant C which may depend on d and the cardinality of F only. However we won't need this stronger result.

Let us draw some consequences of the main theorem (already stated in the introduction).

Corollary 3.5. (*No large torsion balls*) There is an integer $N_2 = N_2(d) \in \mathbb{N}$ such that if F is a finite subset of $SL_d(\mathbb{C})$ containing 1, then either it generates a finite subgroup, or $(F \cup F^{-1})^{N_2(d)}$ contains an element of infinite order. Furthermore if F generates a non virtually nilpotent subgroup, then we can find the element of infinite order already in $F^{N_2(d)}$.

Observe that easy examples in the $\{ax+b\}$ group show that for arbitrary $N \in \mathbb{N}$ we may find a finite set F such that the group generated by F is infinite and virtually abelian, while F^N consists of elements of finite order. For instance, take $F = \{a, tat^{-1}\}$ where a is multiplication by ω (a root of 1 of large order) and t is translation by 1, then the commutator $[a, tat^{-1}]$ is $\neq 1$ and unipotent so of infinite order, while F^k is made of homotheties of ratio ω^k if $k < n$ hence of torsion elements.

Corollary 3.6. There are constants $\eta = \eta(d) > 0$ and $N_1 = N_1(d) \in \mathbb{N}$ such that if F is a finite subset of $SL_d(\overline{\mathbb{Q}})$ generating a non amenable subgroup, then there is a matrix $w \in F^{N_1}$ with an eigenvalue λ such that $h(\lambda) > \eta$.

In particular, if $\mathcal{O}$ is the ring of all algebraic integers, there is an integer $N_1 = N_1(d) \in \mathbb{N}$ such that if F is a finite set of $SL_d(\mathcal{O})$ containing 1, either F generates a virtually solvable subgroup, or there is an archimedean absolute value v on $\overline{\mathbb{Q}}$ extending the canonical absolute value on $\mathbb{Q}$ and a matrix $f \in F^{N_1}$ with at least one eigenvalue of v-absolute value ≥ 2. Observe that this fails for arbitrary finite subsets of $SL_d(\overline{\mathbb{Q}})$. For instance $SL_3(\mathbb{Q}) \cap SO(3, \mathbb{R})$ is dense in $SO(3, \mathbb{R})$ and contains a finitely generated dense subgroup.

4. Preliminary reductions

The goal of this section is to reduce the proof of Theorem 3.1 to the case when $F = \{a, b\}$ is a finite set of two regular semisimple elements generating a Zariski dense subgroup inside $\mathbb{G}(\overline{\mathbb{Q}})$, where $\mathbb{G}$ is a Zariski-connected absolutely simple algebraic group of adjoint type defined over $\overline{\mathbb{Q}}$, and where the underlying vector space is the Lie algebra $\mathfrak{g}$ of $\mathbb{G}$ on which $\mathbb{G}$ acts via the adjoint representation, so that $\mathbb{G} \subset SL(\mathfrak{g})$.

In the setting of Theorem 3.1, we denote by $\mathbb{G}$ the Zariski closure of the group $\langle F \rangle$ generated by F.

4.1. **Reduction to semisimple $\mathbb{G}$.** Let K be a number field and $(e_i)_{1 \leq i \leq d}$ be the canonical basis of $V = K^d$. Let $V = \bigoplus_{1 \leq i \leq m} V_i$ be a direct sum decomposition adapted to this basis, i.e. there are indices $j_1 < ... < j_m$ such that $V_i = span\{e_{j_i}, ..., e_{j_{i+1}-1}\}$. Let P be the group of block upper triangular matrices determined by the corresponding flag, i.e. the parabolic subgroup of GL_d fixing the flag. Let $\rho : P \to GL_d$ be the natural homomorphism that sends a matrix $A = (a_{ij})_{ij} \in P$ to the matrix $\rho(A) = (a'_{ij})_{ij}$ with $a'_{ij} = a_{ij}$ if e_i and e_j belong to the same V_k and $a'_{ij} = 0$ otherwise.

Lemma 4.1. Let $v \in V_K$ be a place of K. Let F be a finite set in $GL_d(K)$. Then

$$E_v(\rho(F)) = E_v(F)$$

Proof. It is straightforward to check that $\|\rho(x)\|_v \leq \|x\|_v$ for every $x \in GL_d(\overline{\mathbb{Q}}_v)$. From this we get the first half of the claimed relation, i.e. $E_v(\rho(F)) \leq \inf_{g\in GL_d(\overline{\mathbb{Q}}_v)} \|\rho(g)\rho(F)\rho(g)^{-1}\| \leq \inf_{g\in GL_d(\overline{\mathbb{Q}}_v)} \|gFg^{-1}\|_v = E_v(F)$.

The second half follows from the remark that $\rho(F)$ can be approximated uniformly by the $\delta F\delta^{-1}$'s for some suitably chosen $\delta \in \Delta(\overline{\mathbb{Q}}_v)$, where Δ is the group of block scalar matrices associated with the V_i's. Indeed we get

$$\begin{aligned} E_v(F) &= \inf_{g\in GL_d(\overline{\mathbb{Q}}_v)} \left\|gFg^{-1}\right\|_v = \inf_{g\in GL_d(\overline{\mathbb{Q}}_v)} \inf_{\delta\in\Delta(\overline{\mathbb{Q}}_v)} \left\|g\delta F\delta^{-1}g^{-1}\right\|_v \\ &\leq \inf_{g\in GL_d(\overline{\mathbb{Q}}_v)} \left\|g\rho(F)g^{-1}\right\|_v = E_v(\rho(F)) \end{aligned}$$

□

Lemma 4.2. In Theorem 3.1, we may assume that $\langle F\rangle$ acts irreducibly on $\overline{\mathbb{Q}}^d$.

Proof. Indeed let $\overline{\mathbb{Q}}^d = \bigoplus_{1\leq i\leq m} V_i$ be such that the associated flag is a decomposition series for $\langle F\rangle$. The map ρ defined above satisfies $e(\rho(F)) = e(F)$ by the lemma above. Moreover $\langle F\rangle$ is virtually solvable if and only if $\rho(\langle F\rangle)$ is virtually solvable and if and only if each $\rho_i(\langle F\rangle)$ is virtually solvable, where ρ_i is the restriction to V_i. As $e(\rho(F)) \geq e(\rho_i(F))$ for each i, we see that it is enough to prove the theorem on each V_i, on which $\rho_i(\langle F\rangle)$ acts irreducibly. □

Hence we may assume that $\mathbb{G}$ acts irreducibly on $\overline{\mathbb{Q}}^d$. But this implies that $\mathbb{G}^\circ$ is a connected reductive group. Indeed, if R_u is the unipotent radical of $\mathbb{G}^\circ$, then R_u, being unipotent, admits a non-zero subspace of fixed points V inside $\overline{\mathbb{Q}}^d$. As R_u is normal in $\mathbb{G}$, V is stabilized by $\mathbb{G}$ itself, hence must be the whole of $\overline{\mathbb{Q}}^d$, by irreducibility. Clearly we may assume that $\mathbb{G}^\circ$ has a non-trivial semisimple part, otherwise $\mathbb{G}$ is virtually solvable. We can now push further and get rid of the non-connectedness of $\mathbb{G}$.

Proposition 4.3. In Theorem 3.1, we may assume that $\mathbb{G}$ is Zariski-connected, semisimple, and acts irreducibly on $\overline{\mathbb{Q}}^d$.

Proof. First we have:

Lemma 4.4. We can assume that $\mathbb{G}^\circ$ is reductive and acts irreducibly on $\overline{\mathbb{Q}}^d$.

Proof. As $\mathbb{G}^\circ$ is reductive, we may find a direct sum decomposition $\overline{\mathbb{Q}}^d = \bigoplus_{1\leq i\leq m} V_i$ into irreducible subspaces of $\mathbb{G}^\circ$. Let ρ_i be the restriction of $\mathbb{G}^\circ$ to V_i. Since $\mathbb{G}$ acts irreducibly, $\mathbb{G}/\mathbb{G}^\circ$ permutes transitively the V_i's. Since $\mathbb{G}^\circ$ is not solvable, we may assume that one of the $\rho_i(\mathbb{G}^\circ)$ is no solvable, say $i = 1$. Let $\mathbb{H}$ be the stabilizer of V_1. Then $[\mathbb{G} : \mathbb{H}] \leq m \leq d$. By Lemma (4.6)

below, we may find a finite set F_0 in $(F \cup \{1\})^{2d-1}$ such that $\langle F_0 \rangle = \mathbb{H} \cap \langle F \rangle$ and is Zariski dense in $\mathbb{H}$. Since $e(F_0) \leq e(F^{2d-1}) \leq (2d-1)e(F)$, we may assume that $F = F_0$ and $\mathbb{G}^\circ$ acts irreducibly. □

Lemma 4.5. Suppose $\mathbb{G}^\circ$ is reductive. Let $\mathbb{S}$ be the semisimple part of $\mathbb{G}^\circ$ ($\mathbb{S} = [\mathbb{G}^\circ, \mathbb{G}^\circ]$) and $\mathcal{Z}$ be the centralizer of $\mathbb{S}$ in $\mathbb{G}$. Then $[\mathbb{G} : \mathcal{Z}\mathbb{S}] \leq c(d)$, where $c(d)$ is a constant depending only on d.

Proof. The group $\mathbb{S}$ is normal in $\mathbb{G}$, let $\sigma : \mathbb{G} \to Aut(\mathbb{S})$ be the map given by conjugation. It induces $\overline{\sigma} : \mathbb{G} \to Out(\mathbb{S})$. But $Out(\mathbb{S})$ is a finite group whose order depends only on the Dynkin diagram of $\mathbb{S}$, hence is bounded in terms of d only (see [4] 14.9). Let $\mathbb{H}$ be the kernel of $\overline{\sigma}$. Then $[\mathbb{G} : \mathbb{H}] \leq c(d)$ by the latter remark. On the other hand, by definition of $\mathbb{H}$, $\mathbb{H} = \mathcal{Z}\mathbb{S}$. □

Lemma 4.6. Let F be a finite subset of a group Γ containing 1. Assume that the elements of F (together with their inverses) generate Γ. Let Γ_0 be a subgroup of index k in Γ. Then F^{2k+1} contains a generating set of Γ_0.

Proof. It is clear that F^k contains a set of representatives for each left coset in Γ/Γ_0, say $\{s_1, ..., s_k\}$. Similarly, $(F^{-1})^k$ contains a set of representatives of the left cosets, say $\{u_1, ..., u_k\}$. Consider all elements of Γ_0 of the form $s_i f u_j^{-1}$ for $i, j \in [1, k]$ and $f \in F$. They all belong to F^{2k+1}. It is straightforward to verify that, together with their inverses, they generate Γ_0. □

The same argument as in Lemma 4.4 shows that we may assume that $\mathbb{G} = \mathcal{Z}\mathbb{S}$. Since $\mathbb{G}^\circ$ acts irreducibly on $\overline{\mathbb{Q}}^d$, and $\mathbb{S}$ is completely reducible, it follows that $\mathbb{S}$ also acts irreducibly. By Shur's lemma, $\mathcal{Z}$ must consist of scalar matrices. But, as $F \subset SL_d(\overline{\mathbb{Q}})$, it must be that $\mathbb{G} \subset SL_d$ also, hence $\mathcal{Z} = \{1\}$, and $\mathbb{G} = \mathbb{S}$ is a Zariski connected semisimple algebraic subgroup of SL_d. □

4.2. **Minimal displacement on symmetric spaces and Bruhat-Tits buildings.** In this paragraph, we give a geometric interpretation of the minimal norm $E_v(F)$ and prove Lemma 4.10, which will be key in the proof of the main theorem. Here k will denote a local field of characteristic 0 and $\mathbb{G}$ a Zariski-connected semisimple k-split k-algebraic subgroup of $SL_d(k)$. Let $\overline{k}$ be a fixed algebraic closure of k and $|\cdot|_k$ the absolute value on $\overline{k}$ induced by the one on k. Let $\mathcal{BT}(\mathbb{G}, k)$ be the Bruhat-Tits building (resp. the symmetric space if k is Archimedean) associated to $\mathbb{G}(k)$. Let $\mathfrak{g}$ be the Lie algebra (over k) of $\mathbb{G}(k)$ and $\mathfrak{a}$ a Cartan subalgebra defined over k.

We first explain how to find a norm on k^d (depending on the way $\mathbb{G}$ sits in SL_d) which will be well suited for our purposes. We first examine the case when k is Archimedean. Then according to a theorem of Mostow (see [20]), there exists a hermitian positive definite scalar product on k^d according to which $\mathbb{G}(k)$ is self-adjoint, i.e. $\mathbb{G}(k)^* = \mathbb{G}(k)$. Up to conjugating $\mathbb{G}(k)$ in

$SL_d(k)$ (equivalently up to choosing a suitable basis of k^d), we may assume that this scalar product is the standard hermitian scalar product $\sum_{i=1}^d x_i\overline{y_i}$. It induces a hermitian scalar product on the Lie algebra $\mathfrak{sl}_d(k)$ defined by $tr(X^*Y)$. Then also $\mathfrak{g}^* = \mathfrak{g}$, and $\mathfrak{a}$ coincides with the diagonal matrices in $\mathfrak{g}$. The restriction of this scalar product to $\mathfrak{g}$ is proportional to the one induced by the Killing form on $\mathfrak{g}$ and the Cartan involution $-X^*$ on $\mathfrak{g}$, i.e. $\langle X, Y\rangle = B_{\mathfrak{g}}(X^*, Y)$. Let $\|\cdot\|_k$ denote both this norm on k^d (resp. $\mathfrak{g}$) and the associated operator norm on $GL_d(k)$ (resp. $GL(\mathfrak{g})$). We denote by K_0 the stabilizer of $\langle X, Y\rangle$ in $\mathbb{G}(k)$ and A the Cartan subgroup with Lie algebra $\mathfrak{a}$. We let x_0 be the base point of $\mathcal{BT}(SL_d, k)$ given by the maximal compact subgroup of $SL_d(k)$ equal to the stabilizer of $\|\cdot\|_k$. By Proposition 4.7, the symmetric space $\mathcal{BT}(\mathbb{G}, k) \simeq \mathbb{G}(k)/K_0$ embeds isometrically as the closed an totally geodesic subspace of $\mathcal{BT}(SL_d, k)$ defined as the $\mathbb{G}(k)$-orbit of x_0 (see also Eberlein [12] 2.6).

Now assume that k is non Archimedean. There is an $\mathcal{O}_k$-lattice L in k^d which is invariant under $\mathbb{G}(\mathcal{O}_k)$: one may choose L to be $(\mathcal{U}_{\mathbb{Z}} \otimes \mathcal{O}_k) \cdot v_+$ where v_+ is a heighest weight vector of $\mathfrak{g} = Lie(\mathbb{G})$ acting on k^d and $\mathcal{U}_{\mathbb{Z}}$ is the integer points of the universal enveloping algebra of $\mathfrak{g}$ (see Steinberg [26] Section 2). From the integer version of the Poincaré-Birkhoff-Witt theorem ([26] Section 2 Theorem 2), the $\mathbb{Z}$-lattice $\mathcal{U}_{\mathbb{Z}} \cdot v_+$ has a basis made of weight vectors of $\mathfrak{a}$. Therefore L has an $\mathcal{O}_k$-basis made of weight vectors for $\mathfrak{a}$, or equivalently for the maximal torus A with Lie algebra $\mathfrak{a}$. Thus after possibly conjugating $\mathbb{G}(k)$ by an element from $GL_d(k)$ we can assume that $L = \mathcal{O}_k^d$, that $K_0 = \mathbb{G}(\mathcal{O}_k) = \mathbb{G}(k) \cap SL_d(\mathcal{O}_k)$ is a maximal compact subgroup of $\mathbb{G}(k)$ and that A consists of diagonal matrices. We then take $\|\cdot\|_k$ to be the standard norm on k^d. Similarly we will denote $\|\cdot\|_k$ the standard norm on the Lie algebra $\mathfrak{sl}_d(k)$, namely the norm associated to the standard lattice $\mathfrak{sl}_d(\mathcal{O}_k)$. By restriction this gives a norm (still denoted $\|\cdot\|_k$) on $\mathfrak{g}$, hence on $\mathfrak{g} \otimes \overline{k}$. With the above choices, the Cartan decomposition for $\mathbb{G}(k)$ is simply the restriction of the corresponding decomposition for $SL_d(k)$, namely: $\mathbb{G}(k) = K_0AK_0$ and Proposition 4.7 below implies that the Bruhat-Tits building $\mathcal{BT}(\mathbb{G}, k)$ with base point K_0 embeds isometrically and $\mathbb{G}(k)$-equivariantly inside the building $\mathcal{BT}(SL_d, k)$ of $SL_d(k)$ with K_0 being identified with $x_0 = Stab_{SL_d(k)}(L) = SL_d(\mathcal{O}_k)$. One may also have appealed to a theorem of Landvogt about functoriality properties of Bruhat-Tits buildings (see [17]).

On $\mathcal{BT}(\mathbb{G}, k)$ we define the distance d to be the standard left invariant distance on $\mathcal{BT}(\mathbb{G}, k)$ with the following normalization: if $a \in A$, then $d(a \cdot x_0, x_0) = \sqrt{\sum_{i=1}^d (\log |a_i|_k)^2}$, where log is the logarithm in base $|\pi_k^{-1}|_k$, with π_k a uniformizer for $\mathcal{O}_k$ if k is non Archimedean, and the standard logarithm if k is Archimedean. When k is non Archimedean and ℓ is a finite extension

of k, then with this normalization, the distance between adjacent vertices on $\mathcal{BT}(\mathbb{G}, \ell)$ is of order 1 and independent of ℓ.

The following proposition was communicated to us by P.E. Caprace [10] (see [9] for background on $CAT(0)$ spaces):

Proposition 4.7. Let k be a local field and $\mathbb{G}$ a semisimple k-isotropic linear algebraic group, with Cartan decomposition $\mathbb{G}(k) = K_0 A K_0$. Assume that $\mathbb{G}(k)$ acts properly by isometries on a complete $CAT(0)$ space X in such a way that semisimple elements of $\mathbb{G}(k)$ acts by semisimple isometries. Assume that K_0 fixes a point p in X which belongs to a flat P stabilized by A. Then the map $gK_0 \mapsto g \cdot p$ induces (up to renormalizing the metric on X) a $\mathbb{G}(k)$-equivariant isometric embedding f from $\mathcal{BT}(\mathbb{G}, k)$ to X.

Proof. Let $G = \mathbb{G}(k)$ and P_0 the A-invariant flat in $\mathcal{BT}(\mathbb{G}, k)$ containing the base point p_0 associated to K_0. We may assume that P is the unique minimal A-invariant flat containing p, so that $\dim(P) = r = rk(\mathbb{G})$ (see [9] Flat Torus Theorem). But the normalizer $N_G(A)$ permutes the A-invariant flats. On the other hand $N_G(A)$ is generated by A and by $N_G(A) \cap K_0$. It follows that $N_G(A)$ stabilizes P. Hence $g \cdot p_0 \mapsto g \cdot p$ induces an $N_G(A)$-equivariant map f between P_0 and P.

Note first that is is enough to show that f is a homothety from P_0 to P. Indeed up to renormalizing the metric in X, we may then assume that f is an isometry from P_0 to P, i.e. $d(a \cdot p, p) = d(a \cdot p_0, p_0)$. But then for any $g, h \in G$, $d(f(g \cdot p_0), f(h \cdot p_0)) = d(h^{-1} g \cdot p, p) = d(a \cdot p, p) = d(g \cdot p_0, h \cdot p_0)$ if $h^{-1} g = k_1 a k_2$.

The fact that $f : P_0 \to P$ is a homothety follows from the rigidity of Euclidean Coxeter group actions. Indeed $N_G(A)$ contains the affine Weyl group as a co-compact subgroup which acts co-compactly by isometries on both P_0 and P. But any such action is isometric to the standard Coxeter representation (cf. [5]). □

We have (see [6], Lemma 4.5):

Lemma 4.8. For any $f, g \in SL_d(k)$ and $x = g^{-1} \cdot x_0 \in \mathcal{BT}(SL_d, k)$, we have

$$\log \left\| g f g^{-1} \right\|_k \leq d(f \cdot x, x) \leq \sqrt{d} \cdot \log \left\| g f g^{-1} \right\|_k$$

Proof. Since d is left invariant, we may assume that $g = 1$. Then we may write $f = k_1 a k_2$ the Cartan decomposition for f. Since x_0 and $\|\cdot\|_k$ are fixed by K_0, we can assume that $f = a$. Then the estimate is obvious from the normalization we chose for d above. □

A consequence of this lemma is that the logarithm of the minimal norm of a finite set F is comparable to the minimal displacement of F on $\mathcal{BT}(SL_d, k)$.

Lemma 4.9. For every $x \in \mathbb{G}(\overline{k})$

$$\|x\|_k \leq \|Ad(x)\|_k \leq \|x\|_k^d$$

Proof. Up to changing k to a finite extension, we may assume that $x \in \mathbb{G}(k)$. Let $\mathbb{G}(k) = K_0 A K_0$ be the associated Cartan decomposition. With the above notation we see that $Ad(K_0)$ stabilizes the norm $\|\cdot\|_k$ on $\mathfrak{g}$. Therefore if $x = k_1 a k_2$, then $\|x\|_k = \|a\|_k = |a_1(x)|_k$ and $\|Ad(x)\|_k = \|Ad(a)\|_k = \frac{|a_1(x)|_k}{|a_d(x)|_k}$. On the other hand $\det(x) = 1$, so $a_1 \cdot ... \cdot a_d = 1$ and as $|a_1|_k \geq ... \geq |a_d|_k$ we must have $|a_1|_k \leq |a_1|_k/|a_d|_k \leq |a_1|_k^d$. We are done. □

Recall that for a finite set F in $SL_d(k)$ we defined the minimal norm $E_k(F) = \inf_{g \in GL_d(\overline{k})} \|gFg^{-1}\|_k$. Similarly define

$$E_k^{Ad}(F) = \inf_{g \in GL_d(\mathfrak{g} \otimes \overline{k})} \left\|gAd(F)g^{-1}\right\|_k .$$

As in [6], 5.4.1., we will use a projection argument and the fact that $\mathcal{BT}(SL_d, k)$ is a $CAT(0)$ space in order to show that the minimal displacement of F is attained on $\mathcal{BT}(\mathbb{G}, k)$. More precisely:

Lemma 4.10. Let k be a local field and $\mathbb{G}$ a Zariski-connected semisimple k-split k-algebraic subgroup of $SL_d(k)$. Let $\|\cdot\|_k$ be the norm defined above. For every finite set $F \in \mathbb{G}(k)$, we have

$$\begin{aligned} E_k(F) &\leq \inf_{g \in \mathbb{G}(\overline{k})} \left\|gFg^{-1}\right\|_k \leq E_k(F)^{\sqrt{d}} \\ E_k^{Ad}(F) &\leq \inf_{g \in \mathbb{G}(\overline{k})} \left\|Ad(gFg^{-1})\right\|_k \leq E_k^{Ad}(F)^{\sqrt{\dim \mathfrak{g}}} \end{aligned}$$

Proof. The left side of the inequalities is obvious from the definition of $E_k(F)$ and $E_k^{Ad}(F)$. For any $\varepsilon > 0$, one can find a finite extension ℓ of k such that $\inf_{g \in \mathbb{G}(\overline{\mathbb{Q}}_v)} \|gFg^{-1}\|_k \leq \inf_{g \in \mathbb{G}(\ell)} \|gFg^{-1}\|_k + \varepsilon$. By Lemma 4.8

(5)
$$\inf_{g \in \mathbb{G}(\ell)} \log \left\|gFg^{-1}\right\|_v \leq \inf_{g \in \mathbb{G}(\ell)} \max_{f \in F} d(fgx_0, gx_0) \leq \inf_{x \in \mathcal{BT}(\mathbb{G}, \ell)} \max_{f \in F} d(fx, x) + c$$

where the log is in base $|\pi_\ell^{-1}|_k$ and c is the maximal distance from any point in $\mathcal{BT}(\mathbb{G}, \ell)$ to the nearest point in the orbit $\mathbb{G}(\ell) \cdot x_0$. Note that this constant c is independent of the choice of ℓ. Since $\mathcal{BT}(SL_d, \ell)$ is a $CAT(0)$ metric space and $\mathcal{BT}(\mathbb{G}, \ell)$ a closed convex subset, for every $x \in \mathcal{BT}(SL_d, \ell)$, one can define the projection $p(x)$ of x on $\mathcal{BT}(\mathbb{G}, \ell)$ to be the (unique) point that realizes the distance from x to $\mathcal{BT}(\mathbb{G}, \ell)$. Then $d(fx, x) \geq d(fp(x), p(x))$ for any $x \in \mathcal{BT}(SL_d, \ell)$. Therefore

(6)
$$\inf_{x \in \mathcal{BT}(\mathbb{G}, \ell)} \max_{f \in F} d(fx, x) = \inf_{x \in \mathcal{BT}(SL_d, \ell)} \max_{f \in F} d(fx, x)$$

Combining (5) with (6) and Lemma 4.8 we get

$$\inf_{g\in\mathbb{G}(\ell)} \log\left\|gFg^{-1}\right\|_k \leq \sqrt{d}\cdot \inf_{g\in SL_d(\ell)} \log\left\|gFg^{-1}\right\|_k + c$$

Hence

$$\inf_{g\in\mathbb{G}(\overline{k})} \left\|gFg^{-1}\right\|_k \leq (|\pi_\ell^{-1}|_k)^c \inf_{g\in SL_d(\ell)} \left\|gFg^{-1}\right\|_k^{\sqrt{d}} + \varepsilon$$

But ℓ can be taken arbitrarily large, so that $|\pi_\ell^{-1}|_k$ can be taken arbitrarily close to 1, and since c was independent of ℓ and ε was arbitrary, we get

$$\inf_{g\in\mathbb{G}(\overline{k})} \left\|gFg^{-1}\right\|_k \leq \inf_{g\in SL_d(\overline{k})} \left\|gFg^{-1}\right\|_k^{\sqrt{d}} = E_k(F)^{\sqrt{d}}$$

This gives the first line of Lemma 4.10. The second line is obtained in exactly the same way; the only thing to check being that $\mathcal{BT}(Ad(\mathbb{G}),k)$ is indeed embedded isometrically as a closed convex subset of $\mathcal{BT}(SL(\mathfrak{g}),k)$ via the $Ad(\mathbb{G}(k))$-equivariant map induced by the inclusion $Ad(\mathbb{G}(k)) \leq SL(\mathfrak{g})(k)$ with base points $Ad(K_0) \leq K_0^{\mathfrak{g}}$, where $K_0^{\mathfrak{g}}$ is the stabilizer of the norm $\|\cdot\|_k$ on $\mathfrak{g}\otimes k$ in $SL(\mathfrak{g})(k)$ defined above. □

4.3. **Reduction to the adjoint representation.** Our goal in this paragraph is to prove the following:

Proposition 4.11. In Theorem 3.1, we may assume that F is a finite subset of $\mathbb{G}(\overline{\mathbb{Q}})$, where $\mathbb{G}$ is a Zariski-connected absolutely simple algebraic group of adjoint type defined over $\overline{\mathbb{Q}}$, viewed via the adjoint representation as an algebraic subgroup of $SL(\mathfrak{g})$, where $\mathfrak{g}$ is the Lie algebra of $\mathbb{G}$.

According to the above Paragraph 4.1, we may assume that the Zariski closure $\mathbb{G}$ of $\langle F\rangle$ in $SL_d(\overline{\mathbb{Q}})$ is a connected semisimple algebraic group $\mathbb{G}$ acting irreducibly on $\overline{\mathbb{Q}}^d$. Let $\mathfrak{g}$ be the Lie algebra of $\mathbb{G}$, viewed as a subalgebra of $\mathfrak{sl}_d$. It is defined over $\overline{\mathbb{Q}}$, hence over some number field K. Choosing a larger K if needed we may also assume that $\mathbb{G}$ is K-split. For each place v in V_K, we consider a norm $\|\cdot\|_v$ on $\mathfrak{g}\otimes K_v$ given as a standard norm corresponding to some choice of a basis. It gives rise to an operator norm on $GL(\mathfrak{g})$. We can thus define $E_v^{Ad}(F)$ as in the last paragraph by $E_v^{Ad}(F) = \inf_{g\in GL(\mathfrak{g}\otimes\overline{\mathbb{Q}}_v)} \|gAd(F)g^{-1}\|_v$, and $e^{Ad}(F)$ the corresponding weighted sum over the set of all places. Note that $E_v^{Ad}(F)$ is independent of the choice of the norm $\|\cdot\|_v$. We have the following key proposition:

Proposition 4.12. For any finite set $F \in \mathbb{G}(K_v)$, we have

$$E_v(F) \leq E_v^{Ad}(F)^{\sqrt{\dim(\mathfrak{g})}} \tag{7}$$

$$E_v^{Ad}(F) \leq E_v(F)^{d\sqrt{d}}$$

As an immediate corollary, we get:

Corollary 4.13. For any $F \in \mathbb{G}(\overline{\mathbb{Q}})$

$$e(F) \leq d \cdot e^{Ad}(F) \leq d^{5/2} \cdot e(F)$$

Proof of Proposition 4.12: This is an application of the results of Paragraph 4.2. Since the quantities involved do not depend on the choice of the norm $\|\cdot\|_v$ used to define them, we may as well take the norm $\|\cdot\|_v = \|\cdot\|_{K_v}$ defined at the beginning of Paragraph 4.2. We have by the second part of Lemma 4.10,

$$E_v(F) \leq \inf_{g \in \mathbb{G}(\overline{\mathbb{Q}}_v)} \left\|gFg^{-1}\right\|_v \leq \inf_{g \in \mathbb{G}(\overline{\mathbb{Q}}_v)} \left\|Ad(gFg^{-1})\right\|_v \leq E_v^{Ad}(F)^{\sqrt{\dim \mathfrak{g}}}$$

Then by Lemma 4.9 and the first part of Lemma 4.10,

$$E_v^{Ad}(F) \leq \inf_{g \in \mathbb{G}(\overline{\mathbb{Q}}_v)} \left\|Ad(gFg^{-1})\right\|_v \leq \inf_{g \in \mathbb{G}(\overline{\mathbb{Q}}_v)} \left\|gFg^{-1}\right\|_v^d \leq E_v(F)^{d\sqrt{d}}$$

We are done.

Proof of Proposition 4.11: According to Proposition 4.12, we may assume that $\mathbb{G} = Ad(\mathbb{G})$ is acting via the adjoint representation on its Lie algebra $\mathfrak{g}$. It remains to verify that we can reduce to a simple factor of $\mathbb{G}$. Recall that $\mathbb{G}$ is the direct product of its simple factors. As the representation space $\mathfrak{g}$ splits into the $\mathbb{G}$-invariant subspaces corresponding to the simple ideals $(\mathfrak{g}_i)_i$ of $\mathfrak{g}$, and as $h^{Ad}(F) \geq h^{Ad_{|\mathfrak{g}_i}}(F)$ for each i, it is enough to prove the theorem for one of the simple factors.

4.4. **Reduction to a 2-element set.** Let us now explain why we may assume that F has 2 elements. First recall that if $\mathbb{G}$ is a connected absolutely almost simple algebraic group over $\overline{\mathbb{Q}}$, a group element $a \in \mathbb{G}(\overline{\mathbb{Q}})$ is said to be regular if $\ker(Ad(a) - 1)$ has the minimal possible dimension (namely equal to the absolute rank of $\mathbb{G}$). For $A_1 \in \mathbb{N}$, we will say that $a \in \mathbb{G}(\overline{\mathbb{Q}})$ is A_1-regular if $\ker(Ad(a) - \omega)$ has minimal possible dimension for every root of unity ω of order at most A_1 (namely dimension 0 if $\omega \neq 1$ and the absolute rank if $\omega = 1$). It is clear that the subset of A_1-regular elements of $\mathbb{G}$ is a non-empty Zariski open subset of $\mathbb{G}$ consisting of semisimple elements.

Further note that for dimension reasons, if T is a maximal torus of $\mathbb{G}$ and Z is a proper Zariski closed subset of $\mathbb{G}$ invariant under conjugation by T, then the Zariski-closure $\widehat{Z}$ of $\{(gag^{-1}, gbg^{-1}) \in \mathbb{G}^2$ with $g \in \mathbb{G}$, $a \in T$ and $b \in Z$, or $a \in Z$ and $b \in T\}$ is a proper algebraic subset of $\mathbb{G} \times \mathbb{G}$.

Since the minimal height $e(F)$ enjoys property (b) from Proposition 2.13, the reduction to a two-element set will obviously follow from the following more general statement, which will be needed in the final argument in the proof of Theorem 3.1.

Proposition 4.14. Let $\mathbb{G}$ be a connected semisimple algebraic subgroup of $GL_d(\overline{\mathbb{Q}})$ with maximal torus T. Let Z be a proper Zariski closed subset of $\mathbb{G}$ invariant under conjugation by T. Then there is an integer $c = c(\mathbb{G}, Z) > 0$ such that if F is a finite subset of $\mathbb{G}(\overline{\mathbb{Q}})$ generating a Zariski-dense subgroup in $\mathbb{G}$, then $(F \cup \{1\})^{c(d)}$ contains two elements a and b which are regular semisimple, generate a Zariski dense subgroup of $\mathbb{G}$, and satisfy $(a, b) \notin \widehat{Z}$. For any given integer $A_1 \in \mathbb{N}$, by allowing c to depend also on A_1, i.e. $c = c(\mathbb{G}, Z, A_1) > 0$, we may further assume that a and b are A_1-regular.

The key ingredient in this proposition is the Eskin-Mozes-Oh escape lemma, Lemma 4.16 below, which is a consequence of a generalized version of Bezout's theorem about the intersection of finitely many algebraic subvarieties (see Zannier's appendix in [23]):

Theorem 4.15 (Generalized Bezout theorem)**.** *Let K be a field, and let $Y_1, \ldots, Y_p$ be pure dimensional algebraic subvarieties of K^n. Denote by $W_1, \ldots, W_q$ the irreducible components of $Y_1 \cap \ldots \cap Y_p$. Then*

$$\sum_{i=1}^{q} deg(W_i) \leq \prod_{j=1}^{p} deg(Y_j).$$

For an algebraic variety X we will denote by $m(X)$ the sum of the degree and the dimension of each of its irreducible components. The following is borrowed from [13], Lemma 3.2.

Lemma 4.16. *(Eskin-Mozes-Oh escape lemma* [13]*) Given an integer $m \geq 1$ there is $N = N(m)$ such that for any field K, any integer $d \geq 1$, any K–algebraic subvariety X in $GL_d(K)$ with $m(X) \leq m$ and any subset $F \subset GL_d(K)$ which contains the identity and generates a subgroup which is not contained in $X(K)$, we have $F^N \nsubseteq X(K)$.*

In order to apply this lemma to prove Proposition 4.14, we need:

Proposition 4.17. Let $\mathbb{G}$ be a connected semisimple algebraic group over $\mathbb{C}$. There is a proper algebraic subvariety X of $\mathbb{G} \times \mathbb{G}$ such that any pair $(x, y) \notin X$ is made of regular semisimple elements which generate a Zariski-dense subgroup of $\mathbb{G}$.

Proof. Recall the well-known:

Lemma 4.18. The set U of regular semisimple elements of $\mathbb{G}$ is a non-empty Zariski-open subset of $\mathbb{G}$.

Proof. The set U coincides with the set of $g \in \mathbb{G}$ such that $\ker(Ad(g) - 1)$ is of minimal dimension. This is clearly a Zariski-open condition. □

We will make use of Jordan's theorem on finite subgroups of $GL_d(\mathbb{C})$ (see [11]). Recall that according to this theorem, there is a constant $C = C(d) \in \mathbb{N}$, such that if Γ is a finite subgroup of $GL_d(\mathbb{C})$, then Γ contains a abelian subgroup A with $[\Gamma : A] \leq C(d)$. As the kernel of the adjoint representation coincides with the center of $\mathbb{G}$, it follows that the same bound apply for all finite subgroups of $\mathbb{G}(\mathbb{C})$ as long as $\dim(\mathbb{G}) \leq d$. Let $V(\mathbb{G})$ be the proper Zariski-closed subset of $\mathbb{G} \times \mathbb{G}$ consisting of all couples (x, y) such that $[x^{C!}, y^{C!}] = 1$. By Jordan's theorem, if $(x, y) \notin V$, then the subgroup generated by x and y infinite.

Let $(\mathbb{G}_i)_{1 \leq i \leq k}$ be the $\mathbb{C}$-simple factors of $\mathbb{G}$, together with their factor maps $\pi_i : \mathbb{G} \to \mathbb{G}_i$. For convenience, let us denote $\mathbb{G}_0 = \mathbb{G}$. Let X_i, for $0 \leq i \leq k$, be the subset of $\mathbb{G} \times \mathbb{G}$ consisting of couples (x, y) such that the $\mathbb{C}$-subalgebra of $End(\mathfrak{g}_i)$ generated by $Ad(\pi_i(x))$ and $Ad(\pi_i(y))$ is of strictly smaller dimension than the subalgebra generated by the full of $Ad(\mathbb{G}_i)$, where $\mathfrak{g}_i$ is the Lie algebra of $\mathbb{G}_i$. This is a Zariski-closed subset of $\mathbb{G} \times \mathbb{G}$. According to [5] VIII.2 ex.8, each $\mathfrak{g}_i$ is generated by two elements. If follows that X_i is a proper closed subvariety. Also let V_i be the set of couples $(x, y) \in \mathbb{G} \times \mathbb{G}$ such that $(\pi_i(x), \pi_i(y)) \in V(\mathbb{G}_i)$, where $V(\mathbb{G}_i)$ is the proper closed subset defined above.

Finally, let X be the proper closed subvariety $X = U^c \cup \bigcup_i X_i \cup \bigcup_i V_i$. Let us verify that X satisfies the conclusion of the proposition. Suppose $(x, y) \notin X$. Then $(x, y) \in U$ and x,y are regular semisimple. Let $\mathbb{H}$ be the Zariski closure of the group generated by x and y. Let $\mathfrak{h}_i$ be the Lie algebra of $\pi_i(\mathbb{H})$, which is a Lie subalgebra of $\mathfrak{g}_i$. As $\mathfrak{h}_i$ is invariant under $Ad(\pi_i(x))$ and $Ad(\pi_i(y))$, it must be invariant $Ad(\mathbb{G}_i)$, by the assumption that $(x, y) \notin X_i$. Therefore $\mathfrak{h}_i$ is an ideal of $\mathfrak{g}_i$. As $\mathfrak{g}_i$ is a simple Lie algebra, either $\mathfrak{h}_i = \{0\}$ or $\mathfrak{h}_i = \mathfrak{g}_i$. In the former case, this means that $\pi_i(\mathbb{H})$ is finite. However, by assumption $(\pi_i(x), \pi_i(y)) \notin V(\mathbb{G}_i)$, this means that the group generated by $\pi_i(x)$ and $\pi_i(y)$ is infinite. So $\pi_i(\mathbb{H})$ is not finite, $\mathfrak{h}_i = \mathfrak{g}_i$ and $\pi_i(\mathbb{H}) = \mathbb{G}_i$.

On the other hand, since $(x, y) \notin X_0$, the same argument shows that the Lie algebra of $\mathbb{H}$ itself is an ideal in $\mathfrak{g}$. Hence $\mathbb{H}^\circ$ is a normal subgroup of $\mathbb{G}$, hence is the product of the simple factors of $\mathbb{G}$ contained in it. The fact that $\pi_i(\mathbb{H}) = \mathbb{G}_i$ for each i forces $\mathbb{H} = \mathbb{G}$. □

Proof of Proposition 4.14: this is immediate by the combination of Proposition 4.17 and Lemma 4.16.

5. Local estimates on Chevalley groups

5.1. **Notation.** Recall our notation. The group $\mathbb{G}$ is an absolutely simple algebraic group of adjoint type defined over $\overline{\mathbb{Q}}$, viewed via the adjoint representation as an algebraic subgroup of $GL(\mathfrak{g})$, where $\mathfrak{g}$ is the Lie algebra of $\mathbb{G}$. We let L be a number field over which $\mathbb{G}$ splits. The set $F = \{a, b\}$ consists

of two semisimple regular elements of $\mathbb{G}(\overline{\mathbb{Q}})$ which generate a Zariski-dense subgroup of $\mathbb{G}$.

Let T be the unique maximal torus of $\mathbb{G}$ containing a. Let $\Phi = \Phi(\mathbb{G}, T)$ be the set of roots of $\mathbb{G}$ with respect to T. Let r be the absolute rank of $\mathbb{G}$. Let us also choose a Borel subgroup B of $\mathbb{G}$ containing T, thus defining the set of positive roots Φ^+ and a base Π for Φ. For $\alpha \in \Phi$, let $\mathfrak{g}_\alpha$ be the root subspace corresponding to α and $\mathfrak{t} = \mathfrak{g}_0$ be the Lie algebra of T, so that we have the direct sum decomposition

$$\mathfrak{g} = \mathfrak{t} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \tag{8}$$

Let $(\alpha_1, ..., \alpha_r)$ be an enumeration of the base associated to the choice of B. The chosen enumeration of the elements of the base induces a total order on the set of roots, namely two roots $\alpha = \sum n_i \alpha_i$ and $\beta = \sum m_i \alpha_i$ satisfy $\alpha \geq \beta$ iff $(n_1, ..., n_r) \geq (m_1, ..., m_r)$ for the canonical lexicographical order on r-tuples. We may label the roots in decreasing order, so that $\alpha(1) > ... > \alpha(|\Phi^+|) > 0 > \alpha(|\Phi^+| + r + 1) > ... > \alpha(|\Phi| + r)$ is the full list of all roots. Note that $d = \dim \mathfrak{g} = |\Phi| + r$ and that $\alpha(|\Phi^+| + r + i) = -\alpha(|\Phi^+| + 1 - i)$ for $1 \leq i \leq |\Phi^+|$. Also set $\alpha(0) = 0$ and $\alpha(i) = 0$ if $i \in I_r = [|\Phi^+| + 1, |\Phi^+| + r]$. Also for any root α, let i_α be the index such that $\alpha(i_\alpha) = \alpha$.

For every $\alpha \in \Phi^+ \cup \{0\}$, let $\mathfrak{u}_\alpha$ be the subspace of $\mathfrak{g}$ generated by the $\mathfrak{g}_\beta$'s for all roots $\beta > \alpha$.

Lemma 5.1. For each $\alpha \in \Phi^+$, $\mathfrak{u}_\alpha$ is an ideal in $\mathfrak{b} = \mathfrak{t} \oplus \bigoplus_{\alpha \in \Phi^+} \mathfrak{g}_\alpha$. Moreover the sequence of $\mathfrak{u}_\alpha$'s for $\alpha \in \Phi^+$ is a decreasing (with α) sequence of non-trivial ideals in $\mathfrak{b}$ starting with $\mathfrak{u}_0 = \bigoplus_{\alpha \in \Phi^+} \mathfrak{g}_\alpha$, each one being of codimension 1 inside the previous one.

Proof. We have $\mathfrak{u}_\alpha = \bigoplus_{\beta > \alpha} \mathfrak{g}_\beta$. Moreover $[\mathfrak{g}_\gamma, \mathfrak{g}_\beta] \leq \mathfrak{g}_{\gamma+\beta}$ and $\gamma + \beta > \alpha$ for any $\gamma \in \Phi^+ \cup \{0\}$, and so clearly $[\mathfrak{b}, \mathfrak{u}_\alpha] \leq \mathfrak{u}_\alpha$. The second assertion follows from the fact that each $\mathfrak{g}_\alpha$, $\alpha \in \Phi$, has dimension 1. □

We also denote by U_α the unipotent algebraic subgroup of $\mathbb{G}$ whose Lie algebra is $\mathfrak{u}_\alpha$, and U_0 the maximal unipotent subgroup, whose Lie algebra is $\mathfrak{u}_0$. Also for each $\alpha \in \Phi$, we denote by $e_\alpha : \mathbb{G}_a \to \mathbb{G}$ the morphism of algebraic groups corresponding to $X_\alpha \in \mathfrak{g}_\alpha$, i.e. $e_\alpha(t) = \exp(tX_\alpha)$. Recall that $U_\alpha = \prod_{\beta > \alpha} e_\beta(\mathbb{G}_a)$, so any element in U_α can be written as a product of $e_\beta(t_\beta)$'s for $\beta > \alpha$.

Recall that since $\mathfrak{g}$ is a simple Lie algebra, it has a Chevalley basis (canonical up to automorphisms of $\mathfrak{g}$) $\{H_\alpha, \alpha \in \Pi\} \cup \{X_\alpha, \alpha \in \Phi\}$ with $H_\alpha \in \mathfrak{t}$ and $X_\alpha \in \mathfrak{g}_\alpha$. Let $(\omega_\alpha)_{\alpha \in \Pi}$ be the basis of $\mathfrak{t}$ which is dual to Π. Equivalently $\beta(\omega_\alpha) = \delta_{\alpha\beta}$. Then $\{\omega_\alpha, \alpha \in \Pi\} \cup \{X_\alpha, \alpha \in \Phi\}$ is also a basis of $\mathfrak{g}$ and defines a $\mathbb{Z}$-structure $\mathfrak{g}_\mathbb{Z}$ on $\mathfrak{g}$ with $[\mathfrak{g}_\mathbb{Z}, \mathfrak{g}_\mathbb{Z}] \subset \mathfrak{g}_\mathbb{Z}$ (see [26]). Hence for any

field k, we can define $\mathfrak{g}_k = \mathfrak{g}_{\mathbb{Z}} \otimes_{\mathbb{Z}} k$. If K is a number field and v a place of K with corresponding embedding $\sigma_v : K \to K_v$ where K_v is the associated completion of K, then we will use the notation $\mathfrak{g}_v$ to mean $\mathfrak{g}_{K_v}$.

Since the definition of $e(F)$ does not depend on the choice of the basis of $\mathfrak{g}$ used to define the standard norm appearing in the quantities $E_v(F)$, we may as well fix the basis of $\mathfrak{g}$ to be the basis $\{\omega_\alpha, \alpha \in \Pi\} \cup \{X_\alpha, \alpha \in \Phi\}$, which we denote $(Y_1, ..., Y_d)$ with $Y_i = X_{\alpha(i)} \in \mathfrak{g}_{\alpha(i)}$ if $i \notin I_r = [|\Phi^+| + 1, |\Phi^+| + r]$ and $Y_i \in \{\omega_\alpha, \alpha \in \Pi\}$ if $i \in I_r$.

Let $B(X, Y)$ be the Killing form on $\mathfrak{g}$. We have $B(Y_i, Y_j) \in \mathbb{Z}$ for all i, j. Let $\tau : \mathfrak{g} \to \mathfrak{g}$ be the linear defined by $Y_i^\tau = -Y_i$ for $i \in I_r$ and and $X_\alpha^\tau = -X_{-\alpha}$ for each $\alpha \in \Phi$. Then τ is an automorphism of $\mathfrak{g}$ which perserves $\mathfrak{g}_{\mathbb{Z}}$. We set $\phi(X, Y) = -B(X^\tau, Y)$.

We now describe how to choose the norm $\|\cdot\|_v$ on $\mathfrak{g}_v$. First consider the case when v is Archimedean, i.e. $\overline{\mathbb{Q}_v} = \mathbb{C}$. We set $\langle X, Y\rangle_v = \phi(X, \overline{Y})$, and thus get a positive definite scalar product on $\mathfrak{g}_v$ and a norm $\|\cdot\|_v$ on $\mathfrak{g}_v$. Let $\mathbf{K}_v = \{g \in \mathbb{G}(\mathbb{C}), g^\tau = g\}$, where we denoted again by τ the automorphism of $\mathbb{G}(\mathbb{C})$ induced by the Chevalley involution τ. Then $\mathbf{K}_v$ is a maximal compact subgroup of $\mathbb{G}(\mathbb{C})$ and this group coincides with the stabilizer of $\langle\cdot,\cdot\rangle_v$ in $\mathbb{G}(\mathbb{C})$, which in turn coincides with $\{g \in \mathbb{G}(\mathbb{C}), \|Ad(g)\|_v = 1\}$ where the norm is the operator norm associated to $\langle\cdot,\cdot\rangle_v$. Note that $(Y_1, ..., Y_d)$ however is not orthogonal with respect to $\langle\cdot,\cdot\rangle_v$ but the decomposition (8) is orthogonal. Finally observe that according to the Iwasawa decomposition we may write $\mathbb{G}(\mathbb{C}) = \mathbf{K}_v U_0(\mathbb{C}) T(\mathbb{C})$.

Suppose now that v is non Archimedean. We let $\|\cdot\|_v$ be the norm induced on $\mathfrak{g}_v$ by the basis $(Y_1, ..., Y_d)$, i.e. $\|\sum y_i Y_i\|_v = \max_{1\leq i\leq d} |y_i|_v$. Then we set $\mathbf{K}_v$ to be the stabilizer in $\mathbb{G}(\overline{\mathbb{Q}}_v)$ of $\mathfrak{g}_{\mathcal{O}_v} = \mathfrak{g}_{\mathbb{Z}} \otimes_{\mathbb{Z}} \mathcal{O}_v$, where $\mathcal{O}_v$ is the ring of integers in $\overline{\mathbb{Q}}_v$. In this situation, the Iwasawa decomposition (see [15]) reads $\mathbb{G}(\overline{\mathbb{Q}}_v) = \mathbf{K}_v U_0(\overline{\mathbb{Q}}_v) T(\overline{\mathbb{Q}}_v)$. Recall (see [26] Paragraph 1, Lemma 6) that for any $n \in \mathbb{N}$ and any $\alpha \in \Phi$, $\frac{ad(X_\alpha)^n}{n!}$ fixes $\mathfrak{g}_{\mathbb{Z}}$. Hence $\left\|\frac{ad(X_\alpha)^n}{n!}\right\|_v \leq 1$.

Let $c_v = \sup_{\alpha\in\Phi} \frac{\|ad(X_\alpha)\|_v}{\|X_\alpha\|_v}$ if v is Archimedean and set $c_v = 0$ if v is non Archimedean. Then, for any place v and $x \in \overline{\mathbb{Q}}_v$, the following holds

$$\|Ad(e_\alpha(x))\|_v = \left\|1 + ad(xX_\alpha) + \frac{ad(xX_\alpha)^2}{2!} + ... + \frac{ad(xX_\alpha)^d}{d!}\right\|_v \tag{9}$$

$$\leq e^{c_v} \cdot \max\{1, \|xX_\alpha\|_v\}^d \tag{10}$$

for every $\alpha \in \Phi$, where $d = \dim \mathfrak{g}$.

Finally we observe that we have:

Lemma 5.2. For each root $\alpha \in \Phi$, the norm $|\alpha|_v = \sup_{Y\in\mathfrak{t}_v\setminus\{0\}} \frac{|\alpha(Y)|_v}{\|Y\|_v}$ satisfies $|\alpha|_v = 1$ whenever v is non Archimedean.

To see this first note that it is obvious if $\alpha \in \Pi$ since $\alpha(\omega_\beta) = \delta_{\alpha\beta}$. As every $\alpha \in \Phi$ is a linear combination with integer coefficients of elements from Π, we must have $|\alpha|_v \leq 1$. To show this opposite inequality, observe that $\gcd(\alpha(\omega_\beta), \beta \in \Pi) = 1$. Indeed suppose there were a prime number p such that p divides $\gcd(\alpha(\omega_\beta), \beta \in \Pi)$. Then $\alpha = p\alpha_0$ with $\alpha_0 = \sum_{i=1}^r n_i\alpha_i$ for some $n_i \in \mathbb{Z}$ and $\Pi = \{\alpha_1, ..., \alpha_r\}$. But since Φ is reduced, α belongs to some base of the root system say $\alpha = \alpha'_1, ..., \alpha'_r$ ([5] VI.1.5). Since each α_i is a linear combination with integer coefficients of some α'_i 's, we get that $\alpha_0 \in \mathbb{Z}\alpha$, a contradiction.

5.2. **Some local estimates.** We first work locally at each place v. The aim of this subsection is to record two estimates, namely Propositions (5.5) and (5.6) below.

Let now $(e_i)_{1\leq i\leq d}$ be an orthonormal basis for $\mathfrak{g}_{\mathbb{C}}$ such that for each $1 \leq i \leq d$, $e_i \in \mathfrak{g}_{\alpha(i)}$. Note that if $b \in Ad(B(\mathbb{C}))$, then the matrix of b is upper-triangular in the basis $(e_i)_i$.

Lemma 5.3. Let V be a complex vector space of dimension n endowed with a hermitian scalar product $\langle \cdot, \cdot \rangle$. Let $(e_i)_{1\leq i\leq n}$ be an orthonormal basis of V and assume that $b \in SL(V)$ has an upper triangular matrix in this basis. Then

$$\sum_{i<j} |\langle be_i, e_j\rangle|^2 \leq n \cdot \left(\|b\|^2 - 1\right)$$

Proof. Let $\lambda_1 \geq ... \geq \lambda_n \geq 0$ be the eigenvalues of b^*b. According to Cartan's KAK decomposition, we have $\|b\|^2 = \lambda_1$. We have

$$tr(b^*b) = \sum \lambda_i \leq n \cdot \lambda_1 = n \cdot \|b\|^2$$

On the other hand,

$$tr(b^*b) = \sum_{i,j} |\langle be_i, e_j\rangle|^2 = \sum_{i<j} |\langle be_i, e_j\rangle|^2 + \sum_{1\leq i\leq n} |\mu_i|^2$$

where $\mu_1, ..., \mu_n$ are the eigenvalues of b. But $\frac{1}{n}\sum_{1\leq i\leq n} |\mu_i|^2 \geq (\prod |\mu_i|)^{\frac{2}{n}} = 1$ since $\det(b) = 1$. Hence

$$n \cdot \|b\|^2 \geq tr(b^*b) \geq \sum_{i<j} |\langle be_i, e_j\rangle|^2 + n$$

□

Lemma 5.4. Let v be any place. Let $\alpha \in \Phi^+$, $a \in T(\overline{\mathbb{Q}}_v)$ regular, $v_\alpha \in U_\alpha(\overline{\mathbb{Q}}_v)$ and $n_\alpha = e_\alpha(x)$ for some $x \in \overline{\mathbb{Q}}_v$ and let $b = Ad(n_\alpha a v_\alpha n_\alpha^{-1})$. Then if v is Archimedean

$$\|xX_\alpha\|_v \leq \frac{\sqrt{d \cdot (\|b\|_v^2 - 1)}}{|1 - \alpha(a)|_v |\alpha|_v} \tag{11}$$

while if v is non Archimedean

$$\|xX_\alpha\|_v \leq \frac{\|b\|_v}{|1-\alpha(a)|_v |\alpha|_v} \tag{12}$$

where $|\alpha|_v$ is the norm of α viewed as linear form on $\mathfrak{t}_v$.

Proof. First observe that if $m \in U_\alpha$ and $Y \in \mathfrak{t}_v$, then $Ad(m)Y \in Y + \mathfrak{u}_\alpha$, while if $m = e_\alpha(x)$ for some x, then $Ad(m)Y = Y + x[X_\alpha, Y] = Y - \alpha(Y)xX_\alpha$. Let $Y \in \mathfrak{t}_v$ be arbitrary. We have $n_\alpha a v_\alpha n_\alpha^{-1} = aa^{-1}n_\alpha a n_\alpha^{-1} n_\alpha v_\alpha n_\alpha^{-1} = a \cdot e_\alpha((\alpha(a^{-1})-1)x) \cdot n''$ where $n'' \in U_\alpha$. We then compute:

$$bY \in Y + x(1-\alpha(a))\alpha(Y)X_\alpha + \mathfrak{u}_\alpha \tag{13}$$

Suppose first that v is Archimedean

$$\langle bY, X_\alpha \rangle_v = x(1-\alpha(a))\alpha(Y) \|X_\alpha\|_v^2$$

On the other hand $Y = \sum y_i e_i$ for some $y_i \in \overline{\mathbb{Q}}_v$ all zero except if $i \in I_r = [|\Phi^+|+1, |\Phi^+|+r]$. Using Cauchy-Schwarz, we get:

$$|\langle bY, X_\alpha \rangle_v| \leq \|X_\alpha\|_v \|Y\|_v \sqrt{\sum_{i \in I_r} |\langle be_i, e_{i_\alpha} \rangle_v|^2}$$

But b is upper-triangular in the basis $(e_i)_i$ because $n_\alpha a v_\alpha n_\alpha^{-1}$ belongs to the Borel subgroup $B(\overline{\mathbb{Q}}_v)$. We are in a position to apply Lemma 5.3, which yields:

$$|(1-\alpha(a))\alpha(Y)|_v \cdot \|xX_\alpha\|_v \cdot \|X_\alpha\|_v \leq \|X_\alpha\|_v \cdot \|Y\|_v \cdot \sqrt{d \cdot (\|b\|_v^2 - 1)}$$

As this is true for all $Y \in \mathfrak{t}$, we indeed obtain (11).

Now assume v is non Archimedean, then (13) shows that

$$\|x(1-\alpha(a))\alpha(Y)X_\alpha\|_v \leq \|b\|_v$$

which is what we wanted. □

Proposition 5.5. There are explicitely computable positive constants $(C_i)_{1 \leq i \leq 3}$ depending only on $d = \dim \mathfrak{g}$ and $p = |\Phi^+|$ such that for any $a \in T(\overline{\mathbb{Q}}_v)$ regular and $u \in U_0(\overline{\mathbb{Q}}_v)$, we have

$$\|Ad(u)\|_v \leq C_3 \cdot \left\|Ad(uau^{-1})\right\|_v^{C_1} \cdot \left(\prod_{i=1}^{p} \max\{1, L_i\}\right)^{C_2} \tag{14}$$

where $L_i = (|1-\alpha(i)(a)|_v \cdot |\alpha(i)|_v)^{-1}$. Moreover if v is non Archimedean, then (14) holds with $C_3 = 1$.

Proof. Recall that we may write $u = e_{\alpha(p)}(x_p) \cdot ... \cdot e_{\alpha(1)}(x_1)$, where $p = |\Phi^+|$ and $x_i \in \overline{\mathbb{Q}}_v$ for each i. We want to apply Lemma 5.4 recursively starting with $\alpha = \alpha(p)$ and going up to $\alpha(1)$. For each $\alpha \in \Phi^+$ let $u_\alpha = e_{\alpha(i_\alpha - 1)}(x_{i_\alpha - 1}) \cdot ... \cdot e_{\alpha(1)}(x_1)$ and $n_\alpha = e_\alpha(x_\alpha)$. For each $i \in [1, p]$ we have $u_{\alpha(i+1)} a u_{\alpha(i+1)}^{-1} = n_\alpha u_\alpha a u_\alpha^{-1} n_\alpha^{-1} = n_\alpha a v_\alpha n_\alpha^{-1}$, where $\alpha = \alpha(i)$ $v_\alpha = a^{-1} u_\alpha a u_\alpha^{-1} \in U_\alpha$.

We set $b_{p+1} = Ad(uau^{-1})$ and $b_i = Ad(u_{\alpha(i)} a u_{\alpha(i)}^{-1})$. Lemma 5.4 gives for each $i \in [1, p]$,

$$\left\| x_{\alpha(i)} X_{\alpha(i)} \right\|_v \leq f_v \cdot L_i \cdot \|b_{i+1}\|_v$$

where $f_v = \sqrt{d}$ if v is Archimedean, and $f_v = 1$ otherwise. Since $b_{i+1} = Ad(n_{\alpha(i)}) b_i Ad(n_{\alpha(i)}^{-1})$, we have

$$\|b_i\|_v \leq \|b_{i+1}\|_v \cdot e^{2c_v} \cdot \max\{1, \left\| x_{\alpha(i)} X_{\alpha(i)} \right\|_v\}^{2d}$$

where we have used (9). Hence combining the last two lines:

$$\|b_i\|_v \leq \mu_i \cdot \|b_{i+1}\|_v^{2d+1} \tag{15}$$

where $\mu_i = e^{2c_v} f_v^{2d} \max\{1, L_i\}^{2d}$.

On the other hand $\|Ad(u)\|_v \leq \prod_{\alpha \in \Phi^+} \|Ad(e_\alpha(x_\alpha))\|_v$ and using (9) again we obtain

$$\begin{aligned} \|Ad(u)\|_v &\leq e^{pc_v} \cdot \left(\prod_{\alpha \in \Phi^+} \max\{1, \|x_\alpha X_\alpha\|_v\} \right)^d \\ &\leq e^{pc_v} \cdot f_v^{dp} \cdot \left(\prod_{i=1}^{p} \max\{1, L_i\} \right)^d \cdot \left(\prod_{i=2}^{p+1} \|b_i\|_v \right)^d \end{aligned}$$

It remains to estimate the last term in the right hand side. Recursively from (15), we get

$$\prod_{i=2}^{p+1} \|b_i\|_v \leq \|b\|_v^{\sum_{k=0}^{p-1} (2d+1)^k} \cdot \prod_{i=2}^{p} \prod_{k=i}^{p} \mu_k^{(2d+1)^{k-i}}$$

Hence we do indeed obtain a bound of the desired form. □

The above proposition is useful to bound $\|Ad(u)\|_v$ when $\|Ad(uau^{-1})\|_v$ may be large. We now need an estimate (only when v is Archimedean) when this norm is small. Let L_i be defined as in the previous statement.

Proposition 5.6. Suppose v is Archimedean. Then there are positive constants $D_i = D_i(d, p)$ for $i = 1, 2, 3$ such that for any $u \in U_0(\overline{\mathbb{Q}}_v)$ and $a \in T(\overline{\mathbb{Q}}_v)$ regular with $\log \|Ad(uau^{-1})\|_v \leq 1$, we have

$$\log \|Ad(u)\|_v \leq D_3 \cdot L_v^{D_2} \cdot \left(\log \left\| Ad(uau^{-1}) \right\|_v \right)^{D_1}$$

where $L_v = \prod_{i=1}^{p} \max\{1, L_i(a)_v\}$ and $D_1 < 1$.

Proof. In this proof, by a constant we mean a positive number that depends only on d and p. Observe that there exists $\varepsilon_1 > 0$ such that $\sqrt{x^2-1} \leq 2\sqrt{\log x}$ as soon as $x \geq 1$ and $\log x \leq \varepsilon_1$. We keep the notations of the proof of the previous proposition. Applying Lemma 5.4, we thus obtain that as soon as $\ell_{i+1} \leq \varepsilon_1$

$$\left\|x_{\alpha(i)}X_{\alpha(i)}\right\|_v \leq 2\sqrt{d}\cdot L_i \cdot \sqrt{\ell_{i+1}}$$

where we set $\ell_i = \log \|b_i\|_v$ for each $i \in [1,p]$, and $\ell = \ell_{p+1} = \log \|Ad(uau^{-1})\|_v$.

We may choose a smaller ε_1 so that

$$\|Ad(e_\alpha(x))\|_v \leq 1 + 2c_v \|xX_\alpha\|_v$$

for each $\alpha \in \Phi^+$ as soon as $|x|_v \leq \varepsilon_1$ as we see from (9). Hence if $\sqrt{\ell_{i+1}} \leq \frac{\varepsilon_1}{2\sqrt{d}\cdot L\cdot}$, then

$$\|b_i\|_v \leq \|b_{i+1}\|_v \cdot \left(1 + 4\sqrt{d}\cdot L_i \cdot \sqrt{\ell_{i+1}}\right)^2$$

or

$$\begin{aligned} \ell_i &\leq \ell_{i+1} + 8\sqrt{d}\cdot L_i \cdot \sqrt{\ell_{i+1}} \\ &\leq C\cdot L \cdot \sqrt{\ell_{i+1}} \end{aligned}$$

for some constant C. Applying this recursively, we see that, as soon as L is bigger than some constant, if $\ell \leq \frac{\varepsilon_1^{2^{p+1}}}{L^{3^{p+1}}}$ then, for each $i \in [1,p]$, $\ell_i \leq \frac{\varepsilon_1^{2^{p+1}}}{L^{3^{p+1}}}$ and

$$\ell_i \leq C' \cdot L^2 \cdot \ell^{\frac{1}{2^{p+1-i}}}$$

for each $i \in [1,p]$ and some constant C'. On the other hand $\|Ad(u)\|_v \leq \prod_{\alpha\in\Phi^+} \|Ad(e_\alpha(x_\alpha))\|_v \leq \prod_{\alpha\in\Phi^+} e^{c_v\|x_\alpha X_\alpha\|_v}$ and

$$\begin{aligned} \log\|Ad(u)\|_v &\leq c_v \cdot \sum_{i=1}^{p} \left\|x_{\alpha(i)}X_{\alpha(i)}\right\|_v \leq C''\cdot L\cdot \sqrt{\sum_{2\leq i\leq p+1}\ell_i} \\ &\leq C'''\cdot L^2 \cdot \ell^{\frac{1}{2^{p+1}}} \end{aligned}$$

On the other hand the cruder bound obtained in Proposition 5.5 shows that without a condition on ℓ,

$$\log\|Ad(u)\|_v \leq \log C_3 + C_1\cdot\ell + C_2\cdot\log L$$

hence

$$\log\|Ad(u)\|_v \leq C_0\cdot L$$

for some constant C_0 if $\ell \leq 1$ and L larger than some constant. Take $D_1 = \frac{1}{2^{p+1}}$, $D_2 \geq 1 + \left(\frac{3}{2}\right)^{p+1}$ and $D_3 \geq \max\{\frac{C_0}{\varepsilon_1}, C'''\}$. Then if $\ell \geq \frac{\varepsilon_1^{2^{p+1}}}{L^{3^{p+1}}}$, we have $D_3\cdot L^{D_2}\cdot \ell^{\frac{1}{2^{p+1}}} \geq D_3\cdot L\cdot\varepsilon_1 \geq C_0\cdot L$. Therefore as soon as $\ell \leq 1$ and L larger than some constant say C_4, we have

$$\log\|Ad(u)\|_v \leq D_3\cdot L^{D_2}\cdot \ell^{D_1}$$

Hence up to changing D_3 into $D_3 C_4^{D_2}$ if necessary, we obtain the desired result. □

6. Global bounds on arithmetic heights

Recall our notations. $\mathbb{G}$ was a Chevalley group of adjoint type and T a maximal torus. We had set the basis $(Y_1, ..., Y_d)$ as obtained from a Chevalley basis of $\mathfrak{g} = Lie(\mathbb{G})$, with $Y_i = X_{\alpha(i)}$ if $i \notin I_r$ and $Y_i \in \{\omega_\alpha, \alpha \in \Pi\}$ if $i \in I_r$. Also $\mathfrak{g}_\mathbb{Z}$ denotes the integer lattice generated by the basis $(Y_1, ..., Y_d)$. Recall further that for $X, Y \in \mathfrak{g}$ we had set $\phi(X, Y) = -B(X^\tau, Y)$ where B is the Killing form and τ the Chevalley involution. Note that (8) is an orthogonal decomposition for the bilinear form ϕ.

We will consider the elements $A = Ad(a)$ and $B = Ad(b)$ from $F = \{a, b\} \subset \mathbb{G}(\overline{\mathbb{Q}})$ with $a \in T$ as matrices in the basis $(Y_1, ..., Y_d)$. Then A is diagonal and $B = (b_{ij})_{ij} \in SL_d(\overline{\mathbb{Q}})$. Consider the regular function on $\mathbb{G}$ given by $f(g) = g_{dd}$ in this basis. Observe the following:

- for every $t \in T$, we have $f(tgt^{-1}) = f(g)$.
- for every $t \in T$, $f(t) = \alpha(d)(t)$, hence f is not constant.
- $\phi(Ad(g)Y_d, Y_d) = f(g)\phi(Y_d, Y_d)$,
- for every place v we have $|f(g)|_v \leq ||Ad(g)||_v$.

The goal of this section is to prove:

Proposition 6.1. For every $n \in \mathbb{N}$ and any $\alpha > 0$ there is $\eta > 0$ and $A_1 > 0$ such that if $F = \{a, b\}$ is a subset of $\mathbb{G}(\overline{\mathbb{Q}})$ with $a \in T(\overline{\mathbb{Q}})$ such that $e(F) < \eta$ and $\deg(\alpha_i(a)) > A_1$ for each positive root α_i, then we have for every $i \in \mathbb{N}$, $1 \leq i \leq n$

$$h(f(b^i)) < \alpha$$

where f is the function defined above.

The proof of this proposition makes use of the local estimates obtained in the previous section as well as Bilu's equidistribution theorem (see below). The proof will occupy the next two subsections. First we collect the local estimates and see what bounds they give us. Then we use Bilu's theorem to show that the remainder terms give only a small contribution to the height.

6.1. Preliminary upper bounds.

Recall that the $(C_i)_{1 \leq i \leq 3}$'s and $(D_i)_{1 \leq i \leq 3}$'s are the constants obtained in Propositions 5.5 and 5.6. For $A \geq 1$ and $x \in \overline{\mathbb{Q}}$ we set

$$h_\infty^A(x) := \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_\infty, |x|_v \geq A} n_v \cdot \log^+ |x|_v \tag{16}$$

where the sum is limited to those $v \in V_\infty$ such that $|x|_v \geq A$. In this paragraph, we prove the following.

Proposition 6.2. There are positive constants A_0, C_4, and D_4 such that if $\varepsilon > 0$ and $A \geq A_0$ are arbitrary, then for any $j \in \mathbb{N}$ and any $a, b \in \mathbb{G}(K)$ two regular semisimple K-rational elements for some number field K with $a \in T(K)$,

(17)
$$\frac{h(f(b^j))}{j} \leq \frac{4 \log A}{\varepsilon} e(\{a,b\}) + D_4 A^{D_2} \varepsilon^{D_1} + C_4 \sum_{1 \leq i \leq p} \left(h_f(\delta_i^{-1}) + h_\infty^A(\delta_i^{-1}) \right)$$

where $\delta_i = 1 - \alpha_i(a)$ for each positive root α_i.

We set as before $F = \{a, b\}$. For each place v let $s_v > \log(E_v^{Ad}(F))$ be some real number. According to Lemma 4.10, there exists $g_v \in \mathbb{G}(\overline{\mathbb{Q}}_v)$ such that $\|Ad(g_v F g_v^{-1})\|_v \leq e^{s_v}$. Since $\mathbb{G}(\overline{\mathbb{Q}}_v) = \mathbf{K}_v U_0(\overline{\mathbb{Q}}_v) T(\overline{\mathbb{Q}}_v)$, and $\mathbf{K}_v$ stabilizes the norm, we may assume that $g_v \in U_0(\overline{\mathbb{Q}}_v) T(\overline{\mathbb{Q}}_v)$, i.e. $g_v = u_v \cdot t_v$. Since t commutes with a we get

$$\begin{aligned} \left\| Ad(u_v a u_v^{-1}) \right\|_v &\leq e^{s_v} \\ \left\| Ad(u_v b^{t_v} u_v^{-1}) \right\|_v &\leq e^{s_v} \end{aligned}$$

where $b^{t_v} = t_v b t_v^{-1}$.

According to Proposition (5.5) we have

$$\begin{aligned} (18) \left\| Ad(b^{t_v}) \right\|_v &\leq e^{s_v} \cdot \|Ad(u_v)\|_v^d \\ &\leq e^{s_v} \cdot C_3^d \cdot \left\| Ad(u_v a u_v^{-1}) \right\|_v^{dC_1} \cdot \left(\prod_{i=1}^p \max\{1, L_i(a)_v\} \right)^{dC_2} \\ &\leq C_3^d \cdot \left(\prod_{i=1}^p \max\{1, L_i(a)_v\} \right)^{dC_2} \cdot e^{s_v(1+dC_1)} \end{aligned}$$

with $C_3 = 1$ if v is non Archimedean. Let $L_v = \prod_{i=1}^p \max\{1, L_i(a)_v\}$. We get

$$\log \left\| Ad(b^{t_v}) \right\|_v \leq d \log C_3 + dC_2 \cdot \log L_v + (1 + dC_1) \cdot s_v \tag{19}$$

Now assume v is Archimedean. According to (5.6) we have constants $D_i > 0$ with $D_1 < 1$ such that if $s_v \leq 1$ then

$$\begin{aligned} (20) \quad \log \left\| Ad(b^{t_v}) \right\|_v &\leq s_v + d \log \|Ad(u_v)\|_v \leq s_v + dD_3 L_v^{D_2} \cdot s_v^{D_1} \\ &\leq D_4' L_v^{D_2} \cdot s_v^{D_1} \end{aligned}$$

where $D_4' = dD_3 + 1$. Since $|f(b^j)|_v \leq \|Ad(b^{t_v})\|_v^j$ for each $j \in [1, n]$ and v, we have obtained:

$$\frac{h(f(b^j))}{j} \leq \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_K} n_v \cdot \log \left\| Ad(b^{t_v}) \right\|_v$$

In order to prove Proposition 6.2, we will decompose this sum into four parts. Let $\kappa = \min_i |\alpha_i|_\infty$. We split the set of places v into four parts: $v \in V_\infty$, $s_v \leq \varepsilon$ and $L_v \geq A/\kappa$ (this gives $H_\leq^+$), $v \in V_\infty$, $s_v \leq \varepsilon$ and $L_v < A/\kappa$ (this gives $H_\leq^-$), $v \in V_\infty$ and $s_v > \varepsilon$ (this gives $H_\geq$) and finally $v \in V_f$ (this gives H_f). So

$$\frac{h(f(b^j))}{j} \leq H_\leq^- + H_\leq^+ + H_\geq + H_f$$

Making use of the bound (19) for $H_\leq^+$, $H_\geq$ and H_f and the bound (20) for $H_\leq^-$ respectively, we obtain the following estimates as soon as A is large enough ($\log A > \log A_0 := 1 + dC_1 + d\log C_3 + \log|\kappa|$, we also set $C_4 = 4dC_2$) :

$$H_f \leq (1+dC_1)\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_f} n_v \cdot s_v + (dC_2)\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_f} n_v \cdot \log L_v$$

$$\begin{aligned} H_\geq &\leq \left(\frac{d\log C_3}{\varepsilon} + (1+dC_1)\right)\cdot\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v\geq\varepsilon} n_v\cdot s_v + \frac{C_4}{4}\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v\geq\varepsilon} n_v\cdot\log L_v \\ &\leq \frac{4\log A}{\varepsilon}\cdot\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v>\varepsilon} n_v\cdot s_v + \frac{C_4}{4}\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v>\varepsilon, L_v\geq A/\kappa} n_v\cdot\log L_v \end{aligned}$$

$$\begin{aligned} H_\leq^+ &\leq (2dC_2)\cdot\frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v\leq\varepsilon, L_v\geq A/\kappa} n_v\cdot\log L_v \\ H_\leq^- &\leq \frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_\infty, s_v\leq\varepsilon, L_v<A/\kappa} n_v\cdot D_4' L_v^{D_2}\cdot s_v^{D_1} \leq 2\frac{D_4'}{\kappa^{D_2}}A^{D_2}\varepsilon^{D_1} \leq D_4 A^{D_2}\varepsilon^{D_1} \end{aligned}$$

for $D_4 = 2\frac{D_4'}{\kappa^{D_2}}$, since $n_v \leq 2$ for $v \in V_\infty$.

Note that $e(F) = \frac{1}{[K:\mathbb{Q}]}\sum_{v\in V_K} n_v\cdot s_v$, so the above bounds give :

(21)

$$\frac{h(f(b^j))}{j} \leq 4\frac{\log A}{\varepsilon}e(F) + \frac{C_4}{2}\frac{1}{[K:\mathbb{Q}]}\left(\sum_{v\in V_\infty, L_v\geq A/\kappa} n_v\cdot\log L_v + \sum_{v\in V_f} n_v\cdot\log L_v\right) + D_4 A^{D_2}\varepsilon^{D_1}$$

On the other hand $\log L_v \leq \sum_{1\leq i\leq p}\log^+ L_i(a)_v$ where $L_i(a)_v = |\delta_i^{-1}|_v/|\alpha_i|_v$ and $\delta_i = 1-\alpha_i(a)$.

Clearly if $L_i(a)_v \geq A/\kappa \geq \kappa^{-2}$ then $|\delta_i^{-1}|_v \geq A$ and $L_i(a)_v \leq |\delta_i^{-1}|_v^2$. We get:

$$\begin{aligned} \sum_{v\in V_\infty, L_v \geq A/\kappa} n_v \cdot \log L_v &\leq \sum_{1\leq i\leq p} \sum_{v\in V_\infty, L_i(a)\geq A/\kappa} n_v \cdot \log^+ L_i(a)_v \\ &\leq 2 \cdot \sum_{1\leq i\leq p} \sum_{v\in V_\infty, |\delta_i|_v \leq A^{-1}} n_v \cdot \log^+ |\delta_i^{-1}|_v \end{aligned} \tag{22}$$

Now note that for $v \in V_f$ we have $|\alpha_i|_v = 1$ according to Lemma 5.2. The product formula applied to δ_i gives:

$$\sum_{v\in V_f} n_v \cdot \log^+ L_i(a)_v = \sum_{v\in V_f} n_v \cdot \log^+ |\delta_i^{-1}|_v = [K:\mathbb{Q}] \cdot h_f(\delta_i^{-1}) \tag{23}$$

Hence combining (21) with (22), (23) we obtain (17) and this ends the proof of Proposition 6.2.

6.2. **Bilu's equidistribution theorem.** We are now going to apply Bilu's equidistribution theorem to show that the last term in estimate (17) become very small when both A is large and $e(F)$ is small.

Theorem 6.3. (Bilu's Equidistribution of Small Points [2]) Suppose $(\lambda_n)_{n\geq 1}$ is a sequence of algebraic numbers (i.e. in $\overline{\mathbb{Q}}$) such that $h(\lambda_n) \to 0$ and $\deg(\lambda_n) \to +\infty$ as $n \to +\infty$. Let $\mathcal{O}(\lambda_n)$ be the Galois orbit of λ_n. Then we have the following weak-$*$ convergence of probability measures on $\mathbb{C}$,

$$\frac{1}{\#\mathcal{O}(\lambda_n)} \sum_{x\in\mathcal{O}(\lambda_n)} \delta_x \underset{n\to+\infty}{\to} d\theta \tag{24}$$

where $d\theta$ is the normalized Lebesgue measure on the unit circle $\{z \in \mathbb{C}, |z| = 1\}$.

Let us first draw two consequences of this equidistribution statement:

Lemma 6.4. For every $\alpha > 0$ there is $A_1 > 0$, $\eta_1 > 0$ and $\varepsilon_1 > 0$ with the following property. If $\lambda \in \overline{\mathbb{Q}}$ is such that $h(\lambda) \leq \eta_1$ and $\deg(\lambda) > A_1$ then

$$h_\infty^{\varepsilon_1^{-1}}(\frac{1}{1-\lambda}) \leq \alpha \tag{25}$$

where $h_\infty^{\varepsilon_1^{-1}}$ was defined in (16).

Proof. We have

$$h_\infty(\frac{1}{1-\lambda}) \leq h(\frac{1}{1-\lambda}) = h(1-\lambda) \leq h_f(\lambda) + h_\infty(1-\lambda) \leq h(\lambda) + h_\infty(1-\lambda)$$

Hence

$$\frac{1}{\deg(\lambda)} \sum_{x \in \mathcal{O}(\lambda)} \log \frac{1}{|1-x|} = h_\infty(\frac{1}{1-\lambda}) - h_\infty(1-\lambda) \leq h(\lambda)$$

and

(26)

$$h_\infty^{\varepsilon_1^{-1}}(\frac{1}{1-\lambda}) = \frac{1}{\deg(\lambda)} \sum_{|1-x| \leq \varepsilon_1} \log \frac{1}{|1-x|} \leq h(\lambda) + \frac{1}{\deg(\lambda)} \sum_{|1-x| > \varepsilon_1} \log |1-x|$$

Consider the function $f_{\varepsilon_1}(z) = \mathbf{1}_{|z-1|>\varepsilon_1} \log|1-z|$. It is locally bounded on $\mathbb{C}$. By Theorem 6.3, for every $\varepsilon_1 > 0$, there must exists $\eta_1 > 0$ and $A_1 > 0$ such that, if $h(\lambda) \leq \eta_1$, and $d = \deg(\lambda) > A_1$, then

$$\left| \frac{1}{\deg(\lambda)} \sum_x f_{\varepsilon_1}(x) - \int_0^1 f_{\varepsilon_1}(e^{2\pi i\theta}) d\theta \right| \leq \frac{\alpha}{3}$$

On the other hand we verify that $\int_0^1 \log|1-e^{2\pi i\theta}| d\theta = 0$. Hence we can choose $\varepsilon_1 > 0$ small enough so that $\left|\int_0^1 f_{\varepsilon_1}(e^{2\pi i\theta}) d\theta\right| \leq \frac{\alpha}{3}$. Combining these inequalities with (26) and choosing $\eta_1 \leq \frac{\alpha}{3}$, we get (25).

□

Lemma 6.5. For every $\alpha > 0$ there exists $\eta > 0$ and $A_1 > 0$ such that for any $\lambda \in \overline{\mathbb{Q}}$, if $h(\lambda) \leq \eta$ and $d = \deg(\lambda) > A_1$, then

$$\left| \frac{1}{\deg(\lambda)} \sum_{v \in V_\infty} n_v \cdot \log|1-\lambda|_v \right| \leq \alpha$$

Proof. The previous lemma shows that the convergence (24) not only holds for compactly supported functions on $\mathbb{C}$, but also for functions with logarithmic singularities at 1. In particular it holds for the function $f(z) = \log|1-z|$, which is exactly what we need, since we check easily that $\int_0^1 f(e^{2\pi i\theta}) d\theta = 0$. □

As a consequence we obtain:

Lemma 6.6. For every $\alpha > 0$ there exists $\eta_0 > 0$ and $A_1 > 0$ such that for any $\lambda \in \overline{\mathbb{Q}}$, if $h(\lambda) \leq \eta_0$ and $d = \deg(\lambda) > A_1$, then

$$h_f(\frac{1}{1-\lambda}) \leq 2\alpha$$

Proof. We apply the product formula to $\delta = 1-\lambda$, which takes the form $h(\delta) = h(\delta^{-1})$, hence

$$h_f(\delta^{-1}) = h_\infty(\delta) - h_\infty(\delta^{-1}) + h_f(\delta)$$

But $h_f(\delta) = h_f(1-\lambda) \leq h_f(\lambda) \leq \eta_0$ and $h_\infty(\delta) - h_\infty(\delta^{-1}) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in V_\infty} n_v \cdot \log |\delta|_v$, which is bounded by α according to Lemma 6.5. We are done. $\square$

The outcome of all this is that each of the terms $h_f(\delta_i^{-1}) + h_\infty^A(\delta_i^{-1})$ in (17) becomes small as soon as $e(F)$ (hence $h(\alpha_i(a))$) becomes small and A becomes large.

6.3. **Proof of Proposition 6.1.** Let $n \in \mathbb{N}$ and $\alpha > 0$ be arbitrary. Let $j \in [1, n]$ an integer and $F = \{a, b\} \subset \mathbb{G}(\overline{\mathbb{Q}})$ with $a \in T(\overline{\mathbb{Q}})$. Then for any $\varepsilon > 0$ and $A > 0$ large enough we obtained the upper bound (17) above. On the other hand we had $h(\alpha_i(a)) \leq d^{3/2} \cdot e(F)$ for each positive root α_i and $\delta_i = 1 - \alpha_i(a)$. Let ε_1, A_1 and η_0 be the quantities obtained in the previous paragraph in Lemmas 6.4 and 6.6. Choose A so that $A^{-1} < \varepsilon_1$ and $A \geq A_0$ and consider (17). Assume that for each $i \in \{1, ..., p\}$ $\deg(\alpha_i(a)) > A_1$. Then Lemmas 6.4 and 6.6 will hold with $\lambda = \alpha_i(a)$ as soon as $e(F) < \eta_0/d^{3/2}$. Hence for each $i = 1, ..., p$

$$\left| h_f(\delta_i^{-1}) + h_\infty^A(\delta_i^{-1}) \right| \leq 2\alpha$$

and

$$\frac{h(f(b^j))}{j} \leq \frac{4 \log A}{\varepsilon} e(\{a, b\}) + D_4 A^{D_2} \varepsilon^{D_1} + 2p(4dC_2)\alpha$$

Now choose $\varepsilon > 0$ so that $2D_4 A^{D_2} \varepsilon^{D_1} < \alpha$. Then choose $\eta > 0$ so that $4\frac{\log A}{\varepsilon}\eta < \alpha$ and $\eta < \eta_0/d^{3/2}$. From (17), we then obtain that if $e(F) < \eta$ and $j \in \mathbb{N}$

$$\frac{1}{j} h(f(b^j)) \leq (2 + p(4dC_2))\alpha$$

Since α was arbitrary we obtain the desired bound.

7. Proof of the statements of Section 3

7.1. **Proof of Theorem 3.1.** The proof Theorem 3.1 will rely Zhang's theorem on small points of algebraic tori. Let $\mathbb{G}_m$ be the multiplicative group and $n \in \mathbb{N}$. On the $\overline{\mathbb{Q}}$-points of the torus $\mathbb{G}_m^n$ we define a notion of height in the following natural way. If $\mathbf{x} = (x_1, ..., x_n) \in \mathbb{G}_m^n$ then $h(\mathbf{x}) := h(x_1) + ... + h(x_n)$ where $h(x_i)$ is the standard height we have been using so far.

Theorem 7.1. (Zhang [32]) Let V be a proper closed algebraic subvariety of $\mathbb{G}_m^n$ defined over $\overline{\mathbb{Q}}$. Then there is $\varepsilon > 0$ such that the Zariski closure V_ε of the set $\{\mathbf{x} \in V,\ h(\mathbf{x}) < \varepsilon\}$ consists of a finite union of torsion coset tori, i.e. subsets of the forms $\boldsymbol{\zeta} H$, where $\boldsymbol{\zeta} = (\zeta_1, ..., \zeta_n)$ is a torsion point and H is a subtorus of $\mathbb{G}_m^n$.

We will need the following lemma, where $\mathbb{G}$ is a semisimple algebraic group over an algebraically closed field, T a maximal torus together with a choice of simple roots Π, and f is the regular function defined at the beginning of the last section:

Lemma 7.2. For every $k \in \mathbb{N}$, the regular functions $f_1, ..., f_k$ defined on $\mathbb{G}$ by $f_i(g) = f(g^i)$ are multiplicatively independent. Namely, if for each i, n_i and m_i are non-negative integers and the f_i's satisfy an equation of the form $f_1^{n_1} \cdot ... \cdot f_k^{n_k} = f_1^{m_1} \cdot ... \cdot f_k^{m_k}$ then $n_i = m_i$ for each i.

Proof. To prove this lemma it is enough to show that for each i one can find a group element $g \in \mathbb{G}$ such that $f_i(g) = 0$ while all other $f_j(g)$'s are non zero. Let H be the copy of PGL_2 corresponding to the roots $\alpha = \alpha(d)$ and $-\alpha = \alpha(1)$ with Lie algebra $\mathfrak{h}$ generated by X_α, $X_{-\alpha}$ and H_α. Clearly it is enough to prove the lemma for the restriction of f to H. Therefore without loss of generality we may assume that $\mathbb{G} = PGL_2$, hence $f(g) = a^2$ if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PGL_2$. Let for instance $D_\lambda = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \in PGL_2$ and $P = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. Set $g_\lambda = PD_\lambda P^{-1}$. Then compute $f(g_\lambda) = 2\lambda - \lambda^{-1}$ and $f_i(g_\lambda) = f(g_{\lambda^i})$. Hence $f_i(g_\lambda) = 0$ if and only if $2\lambda^{2i} = 1$. These conditions are mutually exclusive for distinct values of i. So we are done. □

We can now conclude the proof of Theorem 3.1. According to the reductions made in Section 4 we may assume that $F \subset \mathbb{G}(\overline{\mathbb{Q}})$ where $\mathbb{G}$ is a connected absolutely almost simple algebraic group $\mathbb{G}$ of adjoint type (viewed as embedded in $GL(\mathfrak{g})$ via the adjoint representation), and that the group $\langle F \rangle$ is Zariski dense in $\mathbb{G}$. Let T be a maximal torus in $\mathbb{G}$ and Φ be the corresponding set of roots with set of simple roots Π and let $\alpha(d)$ be the heighest root. The function $f \in \overline{\mathbb{Q}}[\mathbb{G}]$ was defined at the beginning of Section 6 by $f(g) = g_{dd}$ where $\{g_{ij}\}_{1 \le i,j \le d}$ is the matrix of $Ad(g)$ in the Chevalley basis $(Y_1, ..., Y_d)$. Let $f_i(g) = f(g^i)$ and let Ω be the Zariski open subset of $\mathbb{G}$ defined by $\{g,\ f_i(g) \neq 0 \text{ for each } i \le d+1\}$. Let $\mathbf{f}$ be the regular map $\mathbf{f}(g) := (f_1(g), ..., f_{d+1}(g)) : \Omega \to \mathbb{G}_m^{d+1}$. Since $d = \dim \mathbb{G}$, $\operatorname{Im} \mathbf{f}$ is not Zariski dense in $\mathbb{G}_m^{d+1}$. Let V be its Zariski closure. According to the above theorem of Zhang, there is $\mu > 0$ such that the Zariski closure V_μ of $\{\mathbf{x} = (x_1, ..., x_{d+1}) \in V \text{ such that } h(\mathbf{x}) < \mu\}$ is a finite union of torsion coset tori. On the other hand, Lemma 7.2 and the Zariski connectedness of $\mathbb{G}$ shows that V cannot be equal to a finite union of torsion coset tori. Hence V_μ is a proper Zariski closed subset of V. Let $Z_\mu = \Omega^c \cup \mathbf{f}^{-1}\{V_\mu\}$. Then Z_μ is a proper Zariski-closed subset of $\mathbb{G}$. Note that since f is invariant under conjugation by T, Z_μ also is invariant under conjugation by T. Let $\widehat{Z}_\mu$ the Zariski closure of the set $\{(gag^{-1}, gbg^{-1}) \in \mathbb{G}^2 \text{ with } g \in \mathbb{G},\ a \in T \text{ and } b \in Z, \text{ or}$

$a \in Z$ and $b \in T\}$ in $\mathbb{G} \times \mathbb{G}$. Take $n = d+1$ and $\alpha = \mu/n$ in Proposition 6.1, which gives us an $A_1 > 0$ and an $\eta > 0$. According to Proposition 4.14 there is a number $c = c(\mathbb{G}, Z_\mu, A_1) > 0$ such that F^c contains two elements a and b which are A_1-regular semisimple elements, generate a Zariski-dense subgroup of $\mathbb{G}$ and satisfy $(a,b) \notin \widehat{Z}_\mu$. Now let $\varepsilon = \eta/c$ and assume that $e(F) < \varepsilon$. Then $e(\{a,b\}) < \eta$. For some $g \in \mathbb{G}(\overline{\mathbb{Q}})$, $gag^{-1} \in T$, and since h is invariant under conjugation by elements from $\mathbb{G}(\overline{\mathbb{Q}})$, we have $e(\{gag^{-1}, gbg^{-1}\}) < \eta$. We can now apply Proposition 6.1 to see that we must have $h(\mathbf{f}(gbg^{-1})) < \mu$, therefore $gbg^{-1} \in Z_\mu$ and hence $(gag^{-1}, gbg^{-1}) \in \widehat{Z}_\mu$. which gives the desired contradiction. Hence $e(F) > \varepsilon$ and we are done.

7.2. **Proof of Proposition 3.3.** Using the main Theorem 3.1 we see by Proposition 7.4 below that it is enough to prove a weaker form of Proposition 3.3, were we ask for the same estimate together with an additive constant C, i.e. $h(gFg^{-1}) \leq C \cdot \widehat{h}(F) + C$. The proof of this weaker form is independent of the main theorem. It requires however the use of the easier of the two local estimates proved in Section 5, i.e. Proposition 5.5.

Proposition 7.3. Let $G = GL_n(\mathbb{C})$ and B a Borel subgroup of G. For every integers $k, N \geq 2$, let $\mathcal{V}$ be the set of k-tuples $(a_1, ..., a_k) \in G^k$ which leave invariant some finite subset of cardinality at most N in the flag variety G/B. Then $\mathcal{V}$ is a closed algebraic subvariety of G^k.

Proof. We write the proof for $k = 2$. Consider the map $\Phi : G^2 \times (G/B)^N \to (G/B)^{3N}$ which sends $(a, b, x_1, ..., x_n)$ to $((ax_i)_i, (bx_i)_i, (x_i)_i)$. For every two permutations σ and τ of $\{1, ..., N\}$ let $\Delta_{\sigma,\tau}$ be set of $3N$-tuples $((a_i)_i, (b_i)_i, (x_i)_i)$ in $(G/B)^{3N}$ such that $a_i = x_{\sigma i}$ and $b_i = x_{\tau i}$ for each i. It is a closed subvariety of $(G/B)^{3N}$. Let Δ be the union of all $\Delta_{\sigma,\tau}$. Then $\mathcal{V} = \pi_1(\Phi^{-1}(\Delta))$, where π_1 is the projection on G^2, is a closed subvariety of G^2 since G/B is complete. □

Proposition 7.4. Let $G = GL_n(\mathbb{C})$. For every integer k, let $\mathcal{V}$ be the set of k-tuples $(a_1, ..., a_k) \in G^k$ which generate a virtually solvable subgroup. Then $\mathcal{V}$ is a closed algebraic subvariety of G^k.

Proof. Note that for every integer $N \in \mathbb{N}$ if a k-tuple $(a_1, ..., a_k) \in G^k$ leaves invariant a subset of size at most N on the flag variety G/B, then $(a_1, ..., a_k)$ generates a virtually solvable group. So by Proposition 7.3, it suffices to show that conversely there exists a fixed $N = N(n)$ such that if $(a_1, ..., a_k)$ generates a virtually solvable group, then $(a_1, ..., a_k)$ leaves invariant a subset of size at most N on the flag variety G/B. To this end let $\mathbb{G}$ be the Zariski closure of the group generated by $(a_1, ..., a_k)$. Observe that, by induction on n, we may assume that $\mathbb{G}$ acts irreducibly on $\mathbb{C}^n$. Since the connected component $\mathbb{G}^0$ is solvable, Borel's fixed point theorem implies that it fixes a

point on G/B. Let $\mathbb{U}$ be the unipotent radical of $\mathbb{G}^0$. If $\mathbb{U}$ is non trivial it must fix pointwise a non trivial subspace of $\mathbb{C}^n$. As $\mathbb{G}$ normalizes $\mathbb{U}$, $\mathbb{G}$ also must fix that subspace, which contradicts the assumption of irreducibility. Hence $\mathbb{U}$ is trivial and $\mathbb{G}^0$ is a torus. Therefore $\mathbb{G}$ is contained in the normalizer $N(\mathbb{G}^0)$ and $N(\mathbb{G}^0)/Z(\mathbb{G}^0)$ embeds in the Weyl group of G hence has size at most $n!$. We may thus assume that $\mathbb{G}$ centralizes $\mathbb{G}^0$. As we may again assume that $\mathbb{G}$ acts irreducibly, this forces $\mathbb{G}^0$ to be trivial. Hence we are left with the case when $\mathbb{G}$ is finite and we invoke Jordan's theorem to conclude. □

We now prove the weaker form of Proposition 3.3.

Reduction to the adjoint representation.

Let $\mathfrak{g}$ the Lie algebra of $\mathbb{G}$ and $(Y_1, ..., Y_d)$ be the basis of $\mathfrak{g}_{\mathbb{Z}}$ defined in Paragraph 5.1 from a Chevalley basis of $\mathfrak{g}$. Given a local field k, let us call "nice norm" $||\cdot||_{nice,k}$ on $\mathfrak{g}_k$ the norm defined in Paragraph 5.1, that is the Euclidean norm associated to the Killing form on $\mathfrak{g}$ when k is archimedean and the sup norm associated to the basis $(Y_1, ..., Y_d)$ when k is ultrametric. This allows to define the "nice height" $h_{nice}(F)$ by the usual formula (3) where we use the just defined nice norm at each place.

Claim : In proving Proposition 3.3 we may assume that $\mathbb{G} \subseteq SL(\mathfrak{g})$ is absolutely almost simple and acts irreducibly on $\mathfrak{g}_{\overline{\mathbb{Q}}}$ via its adjoint representation and that $h = h_{nice}$.

Thus we assume that $\mathbb{G}$ is a simple algebraic subgroup of $SL_d(\overline{\mathbb{Q}})$. First note that according to Remark 2.11, we may change the basis of $\overline{\mathbb{Q}}^d$, since the heights will be modified only by a bounded additive error. We choose a new basis as follows. Since $\mathbb{G}$ is semisimple, there is a basis say $(v_1, ..., v_d)$ of $\overline{\mathbb{Q}}^d$ defining a new $\mathbb{Q}$-structure for which $\mathbb{G}$ is defined over $\mathbb{Q}$ and $\mathbb{Q}$-split, hence a Chevalley group, and given a maximal $\mathbb{Q}$-split torus T, there exists an integer lattice Λ made of weight vectors which is invariant under $\mathbb{G}(\mathbb{Z})$ (see Steinberg [26] Section 2 Theorem 2). Up to changing the basis again, we may thus assume that $\mathbb{G}$ is defined over $\mathbb{Q}$ and $\mathbb{Q}$-split in SL_d and we have $\mathbb{G}(\mathbb{Z}) = SL_d(\mathbb{Z}) \cap \mathbb{G}(\overline{\mathbb{Q}})$ and $\mathfrak{g}_{\mathbb{Z}} = \mathfrak{g} \cap M_{d,d}(\mathbb{Z})$. Let us denote by ρ this faithful $\mathbb{Q}$-algebraic representation of $\mathbb{G}$ in SL_d.

Lemma 7.5. There are constants $C_0 > 0$ and $L \in \mathbb{N}$ such that for every finite set F in $\mathbb{G}(\overline{\mathbb{Q}})$, $h(\rho(F)) \leq L \cdot h_{nice}(F) + C_0$.

Proof. Given an ultrametric local field k let $||\cdot||_k$ be the standard norm on k^d (as defined in 2.1). It induces an operator norm on $M_{d,d}(k)$, coincides with the standard norm on $M_{d,d}(k)$ associated to the basis of elementary matrices. So its restriction to $\mathfrak{g}_k$ coincides with the standard norm associated to a basis

of $\mathfrak{g}_{\mathbb{Z}}$. This norm also induces an operator norm, say $|||\cdot|||_k$ on $End(M_{d,d}(k))$ (resp. $|||\cdot|||_{\mathfrak{g},k}$ on $End(\mathfrak{g}_k)$). Note that for $g \in \mathbb{G}(k)$, $||g||_{nice,k} = |||Ad(g)|||_{\mathfrak{g},k}$.

We claim that there is a constant $L = L(\rho) \in \mathbb{N}$ such that if $g \in \mathbb{G}(k)$, then $|||Ad(g)|||_{\mathfrak{g},k} \leq |||Ad(g)|||_k \leq |||Ad(g)|||_{\mathfrak{g},k}^L$. Indeed since the maximal compact subgroup $\mathbf{K}_k = \mathbb{G}(\mathcal{O}_k)$ stabilizes $||\cdot||_k$ and also $Ad(\mathbb{K}_0)$ stabilizes $\mathfrak{g}_{\mathcal{O}_k}$ and $M_{d,d}(\mathcal{O}_k)$, we see by the Cartan decomposition, that we may assume $g \in T$. But $\mathfrak{g}_{\mathbb{Z}}$ has a basis made of weight vectors of T, hence for $t \in T$, $|||Ad(t)|||_{\mathfrak{g},k} = \max_{\alpha\in\Phi} |\alpha(t)|_k \geq \max_{\alpha\in\Pi} |\omega_\alpha(t)|_k$ where the ω_α are the fundamental weights. On the other hand if χ is the heighest weight of the representation ρ, and L the heighest coefficient in the expression of χ as a sum of fundamental weights, then $|||Ad(t)|||_k \leq (\max_{\alpha\in\Pi} |\omega_\alpha(t)|_k)^L \leq |||Ad(t)|||_{\mathfrak{g},k}^L$. So we get what we claimed. On the other hand Lemma 4.9 shows that $||x||_k \leq |||Ad(x)|||_k$ for $x \in SL_d(k)$, hence we get for $g \in \mathbb{G}(k)$ $||\rho(g)||_k \leq ||g||_{nice,k}^L$.

Now if k is archimedean, then let $\mathbf{K}_k$ be the maximal compact subgroup of $\mathbb{G}(k)$ defined in Paragraph 5.1, that is the stabilizer of $||\cdot||_{nice,k}$. Then $\rho(\mathbf{K}_k)$ stabilizes some hermitian norm $||\cdot||_k$ on k^d, which need not coincide with the standard norm $||\cdot||_k$. It is however equivalent to it and for some C_0, $||\cdot||_k \leq e^{C_0}||\cdot||_{new}$. Thus $||\rho(g)||_k \leq e^{C_0}||\rho(g)||_{new}$. On the other hand, reasoning as in the ultrametric case $|||Ad(\rho(g))|||_{new} \leq |||Ad(g)|||_{\mathfrak{g},k}^L = ||g||_{nice,k}^L$. Thus $||\rho(g)||_k \leq e^{C_0} \cdot ||g||_{nice,k}^L$.

Taking logarithms and summing over all places, we obtain the desired inequality on heights. □

Suppose now that we knew Proposition 3.3 in the special case of the adjoint repesentation. Thus by assumption there exists $g \in \mathbb{G}(\overline{\mathbb{Q}})$ such that

$$h_{nice}(Ad(gFg^{-1})) \leq C \cdot e^{Ad}(F) + C$$

where e^{Ad} is the quantity defined in Paragraph 4.3 and $C > 0$ is independent of F. But Lemma 7.5 yields $h(\rho(gFg^{-1})) \leq L \cdot h_{nice}(Ad(gFg^{-1})) + C_0$, and according to Corollary 4.13 $e^{Ad}(F) \leq d^{3/2} \cdot e(F)$, while $e(F) \leq \widehat{h}(F) + |\log c|$ by Proposition 2.12 (b). Hence $h(\rho(gFg^{-1})) \leq LCd^{3/2} \cdot \widehat{h}(F) + Cd^{3/2}|\log c| + C_0$. Q.E.D.

End of the proof of Proposition 3.3.

It follows from the Claim that we may assume that $\langle F \rangle$ is Zariski dense in $\mathbb{G}$ and $\mathbb{G}$ is absolutely almost simple of adjoint type and sits in $SL_d(\mathfrak{g})$, where it acts by the adjoint representation on its Lie algebra $\mathfrak{g}$. Let T be a maximal torus in $\mathbb{G}$ and pick a corresponding basis of $\mathfrak{g}_{\mathbb{Z}}$ made of weight vectors say $(Y_1, ..., Y_d)$ as in Section 5. After possibly conjugating F by an element from $\mathbb{G}(\overline{\mathbb{Q}})$ may also assume that $a \in T$.

We can assume that a and b are regular semisimple elements in $\mathbb{G}$, as this condition is Zariski-open. We make the additional assumption that b is generic with respect to T, namely that for any indices i, j the matrix coefficient B_{ij} of $Ad(b)$ in the basis $(Y_1, ..., Y_d)$ is non zero. The set of such pairs (a, b) is clearly Zariski open and non-empty for dimension reasons and because the adjoint representation is irreducible. Nevertheless the result of the proposition holds when a and b are only assumed to be regular semisimple and not contained in a common proper parabolic subgroup of $\mathbb{G}$ (this is a larger Zariski-open subset of $\mathbb{G} \times \mathbb{G}$). This however requires an argument involving the combinatorics of the root system in a key way, which is a little more subtle than the one we are about to give. We will not need this stronger fact.

Let $S \subset [1, d]$ be the set of indices corresponding to the simple roots. So $|S| = rk(\mathbb{G})$. Let $I_r \subset [1, d]$ be the set of indices corresponding to the Y_i's that belong to $\mathfrak{t} = Lie(T)$. For each $j \in S$, let's choose some $i_j \in I_r$ such that $B_{i_j j} B_{j i_j} \neq 0$. Then one can choose a unique point $t \in T(\overline{\mathbb{Q}})$ such that $\alpha_j(t)^2 = \frac{B_{i_j j}}{B_{j i_j}}$ for each $j \in S$. As we may we change $\{a, b\}$ into $\{a, tbt^{-1}\}$. Then $B_{i_j j} = B_{j i_j}$ for every $j \in S$. Moreover we know from (18) that for any place v and any real number $s_v > E_v^{Ad}(F)$, there exists $t_v \in T(\overline{\mathbb{Q}_v})$ such that

$$\left\| Ad(b^{t_v}) \right\|_v \leq C_v^d \cdot \left(\prod_{i=1}^{p} \max\{1, L_i(a)_v\} \right)^{dC_2} \cdot e^{s_v(1+dC_1)}$$

where C_1, C_2, C_∞ are positive constants independent of v and $C_v = 1$ if v is non archimedean, while $C_v = C_\infty$ if v is archimedean. Since every matrix coefficient of $Ad(b^{t_v})$ is bounded by $\|Ad(b^{t_v})\|_v$ if v is non archimedean and by a constant multiple of this norm if v is archimedean, up to enlarging C_∞ if necessary we get that the same bound holds for all matrix coefficients of $Ad(b^{t_v})$, i.e.

(27)

$$\log^+ |\alpha(i)\alpha(j)^{-1}(t_v) B_{ij}|_v \leq d \log C_v + dC_2 \sum_{k=1}^{p} \log^+ L_k(a)_v + (1+dC_1) s_v =: r_v(a)$$

Specializing this for $B_{ij} = B_{ji}$ when $j \in S$ and $i = i_j$ and adding, we obtain

$$2 \log^+ |B_{ij}|_v = \log^+ |B_{ij} B_{ji}|_v \leq 2 r_v(a)$$

On the other hand

$$\begin{aligned} \frac{1}{[K : \mathbb{Q}]} \sum_{v \in V_K} n_v \cdot r_v(a) &\leq d \log C_\infty + dC_2 \sum_{k=1}^{p} (h(\delta_k^{-1}) + \log^+ \frac{1}{\kappa}) + (1 + dC_1) e(F) \\ (28) \qquad &\leq C'_\infty + (1 + dC_1 + dpC_2) e(F) \end{aligned}$$

where C'_∞ is another positive constant, $\delta_k = 1 - \alpha(k)(a)$ for $k \in S$, $\kappa = \min_{k\in S} |\alpha(k)|_\infty$ as above, and where we have used $h(\delta_k^{-1}) = h(\delta_k) \leq h(\alpha(k)(a)) + \log 2 \leq e(F) + \log 2$. Hence for $j \in S$ and $i = i_j$,

$$h(B_{ij}) \leq C'_\infty + (1 + dC_1 + dpC_2)e(F) \tag{29}$$

On the other hand, since $i \in I_r$ $\alpha(i) = 1$ and (27) gives

$$\log^+ |\alpha(j)^{\pm 1}(t_v)B_{ij}|_v \leq r_v(a)$$

$$\log^+ |\alpha(j)^{\pm 1}(t_v)|_v \leq r_v(a) + \log^+ \left|\frac{1}{B_{ij}}\right|_v$$

Taking the weighted sum over all places, we get

$$h(\alpha(j)(t_v)_v), h(\alpha(j)^{-1}(t_v)_v) \leq h(\frac{1}{B_{ij}}) + \frac{1}{[K:\mathbb{Q}]} \sum_{v\in V_K} n_v \cdot r_v(a)$$

which, as $h(B_{ij}^{-1}) = h(B_{ij})$ gives from (28) and (29)

$$h(\alpha(j)(t_v)_v), h(\alpha(j)^{-1}(t_v)_v) \leq 2C'_\infty + 2(1 + dC_1 + dpC_2)e(F) \tag{30}$$

Now let α be an arbitrary root, i.e. $\alpha = \prod_{j\in S} \alpha(j)^{n_j}$ for some integers $n_j \in \mathbb{Z}$. Since there are only finitely many possibilities for the n_j's given $\mathbb{G}$, there is a bound, say N, for the possible sums $\sum |n_j|$. Hence (30) gives

$$h(\alpha(t_v)_v) \leq 2NC'_\infty + 2N(1 + dC_1 + dpC_2)e(F)$$

for every root α. Finally, if i and j are arbitrary indices this time, we get from (27) and (28)

$$\begin{aligned} h(B_{ij}) &\leq \frac{1}{[K:\mathbb{Q}]} \sum_{v\in V_K} n_v \cdot r_v(a) + h(\alpha(i)^{-1}(t_v)_v) + h(\alpha(j)(t_v)_v) \\ &\leq (4N+1)C'_\infty + (4N+1)(1 + dC_1 + dpC_2)e(F) \end{aligned}$$

This ends the proof as $h_{nice}(F) \leq h_{nice}(A) + h_{nice}(B) \leq \sum_{ij} h(A_{ij}) + h(B_{ij})$ and $A_{ij} = 0$ for $i \neq j$ while $h(A_{ii}) \leq e(F)$ by Proposition (2.13) (c).

7.3. **Proof of Corollaries.**

Proof of Corollary 3.6. Let

$$\lambda(F) = \frac{1}{[K:\mathbb{Q}]} \sum_{v\in V_K} n_v \log^+ \Lambda_v(F)$$

From Lemma 2.1, we have $\lambda(F^{d^2}) \geq e(F) - d^2|\log c|$. In particular $\lambda(F^{d^2}) \geq \widehat{h}(F) - d^2|\log c|$ and for every $n \in \mathbb{N}$, $\lambda(F^{nd^2}) \geq n\widehat{h}(F) - d^2|\log c|$. The

theorem says that as F generates a non virtually solvable group, $\widehat{h}(F) > \varepsilon$. On the other hand, we clearly have

$$\lambda(F) \leq \sum_{f\in F}\sum_{i=1}^{d} h(\lambda_i(f)) \leq d \cdot Card(F) \cdot \max\{h(\lambda), \lambda = \lambda_i(f), f \in F\},$$

where $(\lambda_i(f))_{i=1,...,d}$ are the eigenvalues of f. However, it follows from Paragraph 4.4 and in particular Theorem 4.14 that $F^{k(d)}$ always contains two elements a and b which generate a non-virtually solvable subgroup, for some constant $k = k(d)$ independent of F. Applying the above for $F_0 = \{a, b\} \subset F^{k(d)}$ instead of F, we obtain

$$d \cdot 2^{nd^2} \cdot \max\{h(\lambda), \lambda = \lambda_i(f), f \in F_0^{nd^2}\} \geq \lambda(F_0^{nd^2}) \geq n\widehat{h}(F_0) - d^2|\log c|$$

We can finally apply Theorem 3.1, which guarantees that $\widehat{h}(F_0) > \varepsilon(d)$, hence for some eigenvalue λ of a matrix in $F_0^{nd^2}$, $h(\lambda) \geq \frac{1}{d2^{nd^2}}(n\varepsilon - d^2|\log c|) = \eta(d) > 0$ for $n = 2d^2|\log c|/\varepsilon + 1$ for instance. □

Proof of Corollary 3.5. We first assume that the group Γ generated by F is not virtually solvable. Let $\{0\} \leq V_1 \leq ... \leq V_k = \mathbb{C}^d$ be a composition series for Γ. For some index i_0, the composition factor $W = V_{i_0}/V_{i_0+1}$ is irreducible and the image of Γ is $GL(W)$ is not virtually solvable. Therefore we may assume that $\Gamma \leq GL(W)$ and acts absolutely irreducibly on W with $\dim W \geq 2$. According to Burnside's theorem the $\mathbb{C}$-subalgebra generated by the elements of Γ is the full algebra $End_{\mathbb{C}}(W)$. Since $D = \dim End_{\mathbb{C}}(W) = (\dim W)^2 \leq d^2$, there exists a linear basis, say $w_1, ..., w_D$ of $End_{\mathbb{C}}(W)$ in F^{d^2} (start with $w_1 = Id$, then multiply by the elements of F one after the other). Since $\{x \mapsto tr(zx)\}_{z\in End_{\mathbb{C}}(W)}$ account for all linear forms on $End_{\mathbb{C}}(W)$, the linear forms $x \mapsto tr(w_i x)$ must be linearly independent, and the matrix $\{tr(w_i w_j)\}_{1\leq i,j\leq D}$ is invertible. Let L be the field generated by $tr(w_i w_j)$ and $tr(f w_i w_j)$ for $f \in F$ and all i, j. We claim that $\Gamma \leq L[w_1, ..., w_D]$. Indeed for each i, and each $f \in F$, write $f w_i = \sum a_{ij} w_j$ for some $a_{ij} \in \mathbb{C}$. Then as $\{tr(w_i w_j)\}_{1\leq i,j\leq D}$ is invertible, the a_{ij} must belong to L. Since $w_1 = 1$, we see that positive words in F lie all in $L[w_1, ..., w_D]$. On the other hand, the Cayley-Hamilton theorem implies that $f^{-1} \in L[f]$ (every eigenvalue of f belongs to the field generated by $tr(f^k)$ $1 \leq k \leq D$). Finally $\Gamma \leq L[w_1, ..., w_D]$ as claimed. The left regular representation of Γ on $L[w_1, ..., w_D]$ gives us a faithful representation of Γ in $GL_D(L)$. If F^{2d^2+1} consists only of torsion elements, the field L, which belongs to the field generated by the $tr(\gamma)$ for $\gamma \in F^{2d^2+1}$, lies $\overline{\mathbb{Q}}$. We are thus reduced to the case when Γ lies in $GL_D(\overline{\mathbb{Q}})$ and by Corollary 3.6 we are done.

Now assume F generates a virtually solvable group. It is well known (see [31]) that there is an integer $n_0 = n_0(d) \in \mathbb{N}$ such that any virtually solvable

subgroup of $SL_d(\mathbb{C})$ contains a subgroup of index at most n_0 which can be conjugated inside the upper-triangular matrices. Applying Lemma 4.6, we may assume without loss of generality that F is made of upper-triangular matrices. Then for every $a, b \in F$, the commutator $[a, b]$ is a unipotent matrix in $SL_d(\mathbb{C})$, hence is either trivial or of infinite order. If one of them has infinite order, we are done. Otherwise this means that the matrices in F commute. But an abelian group generated by torsion elements is finite. We are done.

The argument above works verbatim without the need to take inverses until the point when F is assumed to consist of upper-triangular matrices. Note that if the elements of F are torsion, then their eigenvalues are roots of unity, hence the group generated by F is virtually nilpotent. This completes the proof of the corollary. □

Proof of Corollary 1.4 from the Introduction. If $tr(\gamma)$ is transcendental for some $\gamma \in F^{2d^2+1}$, then γ has a transcendental eigenvalue λ and the second alternative holds. If $tr(\gamma)$ is algebraic for all $\gamma \in F^{2d^2+1}$, then the argument given in the proof of Corollary 3.5 shows that Γ has a representation in $GL_{d^2}(\overline{\mathbb{Q}})$ with non virtually solvable image. So we are reduced to this situation and the claim is obvious by Corollary 3.6. □

Proof of Corollary 1.5 from the Introduction. If F fixes a point in the Bruhat-Tits building X_k of SL_d over a p-adic field k, then F fixes a vertex of X_k (it fixes the vertices of the smallest simplex containing the fixed point). But vertices of X_k are permuted transitively by the action of $GL_d(k)$. If follows that $E_k(F) = 1$. Hence if F fixes a point on each X_k for k non archimedean, then $e_f(F) = 0$. Hence $e_\infty(F) > \varepsilon$. So there must exist an embedding σ of K in $\mathbb{C}$ such that $\log E_{\mathbb{C}}(\sigma(F)) > \varepsilon$. Then by Lemma 4.8, every point of $X_{\mathbb{C}}$ must be moved by at least ε by some element of F. Q.E.D.

Acknowledgments 7.6. I am grateful to P. Sarnak and A. Yafaev from whom I learned about Bilu's theorem. I thank E. Bombieri, J-B. Bost and A. Chambert-Loir for their insights about diophantine geometry. I am grateful to P.E. Caprace for his invaluable help with buildings. I thank G. Chenevier, E. Lindenstrauss and A. Salehi-Golsefidy for our stimulating conversations. Finally I thank T. Gelander without whom I would not have been led to these mathematics.

References

[1] L. Bartholdi, Y. de Cornulier, *Infinite groups with large balls of torsion elements and small entropy*, to appear in Archiv der Mathematik.

[2] Bilu, Y, *Limit distribution of small points on algebraic tori*, Duke Math. J. **89** (1997), no. 3, 465–476.

[3] Bombieri, E., Gubler, W., *Heights in Diophantine geometry*, New Mathematical Monographs, 4. Cambridge University Press, Cambridge, (2006).
[4] Borel, A., *Linear algebraic groups*, Notes taken by Hyman Bass W. A. Benjamin, Inc., New York-Amsterdam 1969
[5] Bourbaki, N. *Groupes et Algèbres de Lie*, Chapitres 4-5-6 and 7-8, Hermann ed.
[6] Breuillard, E, Gelander, T., *Uniform independence in linear groups*, to appear in Invent. Math.
[7] Breuillard, E., *Heights on GL_2 and free subgroups*, preprint December 2007.
[8] Breuillard, E., *A strong Tits alternative*, preprint April 2008.
[9] Bridson M., Haefliger A., *Metric spaces of non-positive curvature*, Springer-Verlag , (1999), vii, 643 p.
[10] P.E. Caprace, personal communication.
[11] C.W. Curtis, I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, (Interscience, New York) (1962).
[12] P. Eberlein, *Geometry of nonpositively curved manifolds*, Chicago Lectures in Math. (1996).
[13] Eskin, Alex; Mozes, Shahar; Oh, Hee, *On uniform exponential growth for linear groups*, Invent. Math. **160** (2005), no. 1, 1–30
[14] Kazhdan, D., Margulis, G., *A proof of Selberg's hypothesis*, Mat. Sb. (N.S.) **75** (117) 1968 163–168
[15] Iwahori, N., Matsumoto, H., *On some Bruhat decomposition and the structure of the Hecke rings of p-adic Chevalley groups*, Inst. Hautes Études Sci. Publ. Math. No. **25** (1965) 5–48.
[16] Jacobson, N., *Lie algebras*, Interscience Dover (1962).
[17] Landvogt, E., *Some functorial properties of the Bruhat-Tits building*, J. Reine Angew. Math. **518** (2000), 213–241.
[18] Lang, S., *Fundamentals of Diophantine geometry*, Springer-Verlag, New York, (1983).
[19] Masser, Wustholz, *Fields of large transcendence degree generated by values of elliptic functions*, Invent. Math. (1983).
[20] Mostow, G. D., *Self-adjoint groups*, Ann. of Math. (2) **62**, (1955). 44–55.
[21] Onishchik, A. L.; Vinberg, È. B. *Lie groups and algebraic groups*, Translated from the Russian and with a preface by D. A. Leites. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, (1990).
[22] M.S. Raghunathan, *Discrete Subgroups of Lie Groups,* Ergebnisse der Mathematik und Ihrer Grenzgebiete. Band **68** (1972).
[23] Schinzel, A. *Polynomials with special regard to reducibility*, With an appendix by Umberto Zannier. Encyclopedia of Mathematics and its Applications, 77. Cambridge University Press, Cambridge, (2000).
[24] Y. Shalom, *Explicit Kazhdan constants for representations of semisimple and arithmetic groups*, Ann. Inst. Fourier, **50** (2000), no. 3, 833–863.
[25] L. Szpiro, E. Ullmo, S. Zhang, *Equirépartition des petits points*, Invent. Math. **127** (1997), 337–347.
[26] Steinberg, R., *Lectures on Chevalley groups*, Notes prepared by John Faulkner and Robert Wilson. Yale University, New Haven, Conn., (1968).
[27] J. Tits, *Free subgroups of Linear groups*, Journal of Algebra **20** (1972), 250-270.
[28] Thurston, W, *Three-dimensional geometry and topology*, Vol. 1. Edited by Silvio Levy. Princeton Mathematical Series, **35**. Princeton University Press, (1997).

[29] Ullmo, E. *Positivité et discrétion des points algébriques des courbes*, Ann. of Math. (2) **147** (1998), no. 1, 167–179.
[30] Wang, H. C., *Topics on totally discontinuous groups*, in Symmetric spaces (Short Courses, Washington Univ., St. Louis, Mo., 1969–1970), pp. 459–487. Pure and Appl. Math., Vol. 8, Dekker, (1972).
[31] Wehrfritz, B., *Infinite linear groups. An account of the group-theoretic properties of infinite groups of matrices*, Ergeb. Mat. Grenz., **76**, Springer-Verlag, (1973).
[32] Zhang, S., *Small points and adelic metrics*, J. Algebraic Geom. 4 (1995), no. **2**, 281–300
[33] Zhang, S-W., *Equidistribution of small points on abelian varieties*, Ann. of Math. (2) **147** (1998), no. 1, 159–165.

Emmanuel Breuillard, Ecole Polytechnique, France
E-mail address: emmanuel.breuillard@math.polytechnique.fr